library(dplyr)

library(lavaan)

library(lavaanPlot)

library(lavaan)

path <- file.choose()

data <- readLines(path)

header <- data[1]

variables <- strsplit(header, "\\^")[[1]]

datiR\_2022 <- read.table(text = paste(data[-1], collapse = "\n"),

header = FALSE, sep = "^",

col.names = variables,

stringsAsFactors = FALSE)

rm(data, header, path, variables)

AVQ <- datiR\_2022

avq\_clean <- AVQ %>%

mutate(across(c(TRAF, PARCH ),

~ ifelse(. == 5, NA, .)))

invert\_then\_normalize <- function(x) {

if (all(is.na(x))) return(rep(NA, length(x)))

inv\_x <- max(x, na.rm = TRUE) - x

rng <- range(inv\_x, na.rm = TRUE)

(inv\_x - rng[1]) / diff(rng)

}

avq\_scaled <- avq\_clean %>%

mutate(across(

c(

COLLEG, FIDU1, FIDU2, FIDU3, COLCO, COLCOM,

FRCMT, PUNMT, SEDMT, VELMT, PULMT, ATTMT, ORAMT, CBMT, INFMT,

FRTRE, PUTRE, PSTRE, VETRE, PULSTR, COMAT, COMORA, COBIG, INFTR,

FCORS, PUNT, POSSE, VELCO, PULIZ, COMOF, COMOR , CPSBI , TRAF, SITE, PARCH, RISEC),

invert\_then\_normalize,

.names = "norm\_inv\_{col}"

))

normalize <- function(x) {

rng <- range(x, na.rm = TRUE)

(x - rng[1]) / diff(rng)

}

avq\_scaled <- avq\_scaled %>%

mutate(across(c(PUNTIFI10, PUNTIFI8, PUNTIFI1),

normalize, .names = "norm\_{col}"))

avq\_scaled$Connection <- rowMeans(avq\_scaled[, c( "norm\_inv\_COLLEG", "norm\_inv\_COLCO", "norm\_inv\_COLCOM")], na.rm = TRUE)

avq\_scaled$IND\_CLEAN <- rowMeans(avq\_scaled[, c("norm\_inv\_PULIZ", "norm\_inv\_PULMT", "norm\_inv\_PULSTR")], na.rm = TRUE)

avq\_scaled$IND\_SPEED <- rowMeans(avq\_scaled[, c("norm\_inv\_VELCO", "norm\_inv\_VELMT", "norm\_inv\_VETRE")], na.rm = TRUE)

avq\_scaled$IND\_PUNCTUALITY <- rowMeans(avq\_scaled[, c("norm\_inv\_PUNT", "norm\_inv\_PUNMT", "norm\_inv\_PUTRE")], na.rm = TRUE)

avq\_scaled$IND\_SEATS <- rowMeans(avq\_scaled[, c("norm\_inv\_POSSE", "norm\_inv\_SEDMT", "norm\_inv\_PSTRE")], na.rm = TRUE)

avq\_scaled$IND\_WAITING <- rowMeans(avq\_scaled[, c("norm\_inv\_COMOF", "norm\_inv\_ATTMT", "norm\_inv\_COMAT")], na.rm = TRUE)

avq\_scaled$IND\_TIME <- rowMeans(avq\_scaled[, c("norm\_inv\_COMOR", "norm\_inv\_ORAMT", "norm\_inv\_COMORA")], na.rm = TRUE)

avq\_scaled$IND\_COST <- rowMeans(avq\_scaled[, c("norm\_inv\_CPSBI", "norm\_inv\_CBMT", "norm\_inv\_COBIG")], na.rm = TRUE)

rename\_vec <- c(

WAITING = "IND\_WAITING",

SPEED = "IND\_SPEED",

SEATS = "IND\_SEATS",

TIME = "IND\_TIME",

COST = "IND\_COST",

PUNCTUALITY = "IND\_PUNCTUALITY",

CLEAN = "IND\_CLEAN",

CONNECTION = "Connection",

CONGESTION = "norm\_inv\_TRAF",

PARKING = "norm\_inv\_PARCH",

TRUST3 = "norm\_PUNTIFI10",

TRUST1 = "norm\_PUNTIFI1",

TRUST2 = "norm\_PUNTIFI8",

AFFORDABILITY1 = "norm\_inv\_RISEC",

AFFORDABILITY2 = "norm\_inv\_SITE"

)

avq\_scaled\_renamed <- avq\_scaled %>%

rename(!!!rename\_vec)

# SEM MODEL (OARC stats, 2025; Lavaan.org, 2025)

model\_new <- '

Perceived\_accessibility =~ WAITING + SPEED + SEATS + TIME + COST + PUNCTUALITY + CLEAN + CONNECTION

Environment =~ CONGESTION + PARKING

Administration =~ TRUST3 + TRUST1 + TRUST2

Affordability =~ AFFORDABILITY1 + AFFORDABILITY2

Perceived\_accessibility ~ Administration + Environment + Affordability

'

fit <- sem(model\_new, data = avq\_scaled\_renamed, missing = "fiml")

summary(fit, standardized = TRUE, fit.measures = TRUE)

fitMeasures(fit)

# Plot (Lishinski, 2024)

lavaanPlot(model = fit,

node\_options = list(shape = "box", fontname = "Helvetica", fontsize = 14),

edge\_options = list(color = "black", lty = 1, lwd = 1.5),

coefs = TRUE,

stand = FALSE)

# CORRELATION (whalley, na)

vars\_cor <- avq\_scaled\_renamed %>% select(WAITING,SPEED ,SEATS ,TIME, COST ,PUNCTUALITY , CLEAN ,CONNECTION, CONGESTION , PARKING, TRUST3 , TRUST1 , TRUST2, AFFORDABILITY1, AFFORDABILITY2

)

corr\_res\_spearman <- psych::corr.test(vars\_cor, use = "pairwise", method = "spearman")

print(corr\_res\_spearman$r)

print(corr\_res\_spearman$p)

# NORMALITY (Hallquist, 2017)

print(psych::describe(vars\_cor)[, c("skew", "kurtosis")])

# CRONBACH ALPHA (RPUBS, na)

alpha\_sets <- list(

Perceived\_accessibility = avq\_scaled\_renamed %>% select(WAITING , SPEED , SEATS , TIME , COST , PUNCTUALITY , CLEAN , CONNECTION),

Affordability = avq\_scaled\_renamed %>% select(AFFORDABILITY2, AFFORDABILITY1 ),

environment = avq\_scaled\_renamed %>% select( CONGESTION, PARKING),

Administration = avq\_scaled\_renamed %>% select(TRUST3, TRUST1, TRUST2)

)

alphas <- lapply(alpha\_sets, psych::alpha)

alpha\_results <- lapply(alphas, function(a) a$total$raw\_alpha)

print(alpha\_results)

# COMPOSITE RELIABILITY E AVE (JORGENSEN, 2015; GOOGLE, 2025)

print(semTools::compRelSEM(fit))

print(semTools::AVE(fit))

#R-Squared (RDRR.IO, na)

r2\_values <- inspect(fit, "r2")

print(r2\_values)

#other measures (STATISTIKA, 2025)

fitMeasures(fit)

cfi <- fitMeasures(fit, "cfi")

df\_model <- fitMeasures(fit, "df")

df\_baseline <- fitMeasures(fit, "baseline.df")

print(cfi)

print(df\_model)

print(df\_baseline)

pcfi <- cfi \* (df\_model / df\_baseline)

print(pcfi)

#VIF (Bhale and Harpreet, 2024)

vars <- c(vars <- c("WAITING", "SPEED", "SEATS", "TIME", "COST", "PUNCTUALITY", "CLEAN", "CONNECTION", "CONGESTION", "PARKING", "TRUST3", "TRUST1", "TRUST2", "AFFORDABILITY1", "AFFORDABILITY2")

)

calc\_all\_vif <- function(data, variables) {

vif\_results <- numeric(length(variables))

names(vif\_results) <- variables

for (var in variables) {

predictors <- setdiff(variables, var)

formula <- as.formula(paste(var, "~", paste(predictors, collapse = " + ")))

lm\_model <- lm(formula, data = data)

r2 <- summary(lm\_model)$r.squared

vif\_results[var] <- 1 / (1 - r2)

}

return(vif\_results)

}

vif\_all <- calc\_all\_vif(avq\_scaled\_renamed, vars)

print(vif\_all)

# List of reference

# BHALE U. A., and Harpreet S. B. 2024. What is mean by Multicollinearity in SEM?

#URL: https://www.researchgate.net/publication/377691806

# GOOGLE. 2025. how to do convergent and discriminant validity check in R?

# URL:<https://groups.google.com/g/lavaan/c/36xE4mTsvQg>

# JORGENSEN T. D. 2015. semTools: Useful Tools for Structural Equation Modeling. #URL:<https://cran.r-project.org/web/packages/semTools/semTools.pdf>

#HALLQUIST. 2017. Best practices in SEM PSY 597.

#URL: <https://psu-psychology.github.io/psy-597-SEM/12_best_practices/best_practices.html>

# LAVAAN:ORG. 2025. Extracting information.

#URL: <https://lavaan.ugent.be/tutorial/inspect.html>

#LISHINSKI A. 2024. Intro to lavaanPlot.

#URL: <https://cran.r-project.org/web/packages/lavaanPlot/vignettes/Intro_to_lavaanPlot.html>

# OARC stats. 2025. Introduction to Structural Equation Modeling (SEM) in R with #URL:lavaan.<https://stats.oarc.ucla.edu/r/seminars/rsem/>

#RDDR. lavInspect: Inspect or extract information from a fitted lavaan object.

#URL: <https://rdrr.io/cran/lavaan/man/lavInspect.html>

#RPUBS. reliability analysis.

URL: <https://rpubs.com/hauselin/reliabilityanalysis>

# STATISTIKA. 2025. How to interpret SEM model fit results in AMOS.

#URL:<https://www.statistischeberatung.de/images/services/How_to_interpret_SEM_model#-Fit_results_in_AMOS.pdf>

# UNIVERSITY OF FLORIDA. Spearman correlation in R.

#URL: <https://users.phhp.ufl.edu/marsiske/r/AnalyzeCorrelateSpearman.html>

# WHALLEY B. Correlation.

#URL: https://benwhalley.github.io/just-enough-r/correlations.html