Factor Analysis Analysis

Nadia Ahmad

2023-04-22

Library

```
library(readr)
library(readxl)
library(tidyverse)
library(corrplot)
library(ggfortify)
library(psych)
library(gplots)
library(ggpubr)
library(magrittr)
library(REdaS)
library(nFactors)
```

Read the dataset

```
travel <- read_excel("factor_analysis_data.xlsx")
head(travel)
```

```
## # A tibble: 6 x 13
##   travel_frequ~1 discr~2 relig~3 check~4 drop_~5 secur~6 board~7 bagga~8 avoid~9
##            <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1              4       1       8      54      50      71      50      51       2
## 2              4       2       2      52      52      51      53      52       2
## 3              3       2       8      NA      NA      NA      NA      NA       2
## 4              4       2       1      51      53      54      52      57       2
## 5              4       2       2      48     100     100     100     100       1
## 6              4       2       8      50      50      50      50      50       2
## # ... with 4 more variables: asked_deplane <dbl>, detained <dbl>,
## #   age_group <dbl>, gender <dbl>, and abbreviated variable names
## #   1: travel_frequency, 2: discrimination_award, 3: religion_representation,
## #   4: checkin_exp, 5: drop_off_exp, 6: security_exp, 7: boarding_exp,
## #   8: baggage_exp, 9: avoid_travel
```

Dataset contains 228 rows and 13 columns which is still messy. Thus, we'll conduct some data preprocessing steps.

DATA PREPROCESSING

```
# CHECK MISSING VALUE----
# Count the missing values by column wise
```

```
print("Count of missing values by column wise")
```

```
## [1] "Count of missing values by column wise"
```

```
sapply(travel, function(x) sum(is.na(x)))
```

```
##        travel_frequency    discrimination_award religion_representation
##                      26                      47                      51
##             checkin_exp           drop_off_exp           security_exp
##                      72                      79                      74
##            boarding_exp            baggage_exp           avoid_travel
##                      75                      77                      88
##           asked_deplane               detained              age_group
##                      89                      89                      74
##                  gender
##                      74
```

```
# Missing value imputation
# Since our data contains 46 missing value, let's impute with mode
# Function to see mode
calc_mode <- function(x){

  # List the distinct / unique values
  distinct_values <- unique(na.omit(x))

  # Count the occurrence of each distinct value
  distinct_tabulate <- tabulate(match(x, distinct_values))

  # Return the value with the highest occurrence
  distinct_values[which.max(distinct_tabulate)]
}

# Impute missing value----
travel_df <- travel %>%
  mutate(across(everything(), ~replace_na(.x, calc_mode(.x))))
```

```
# CONVERT DATA TYPE----
# Convert all variables into integer
# Convert column 2 to 6 to numeric
travel_df[,c(2, 9:13)] <- lapply(travel_df[,c(2, 9:13)], as.integer)
head(travel_df)
```

```
## # A tibble: 6 x 13
##   travel_frequ~1 discr~2 relig~3 check~4 drop_~5 secur~6 board~7 bagga~8 avoid~9
##            <dbl>   <int>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <int>
## 1              4       1       8      54      50      71      50      51       2
## 2              4       2       2      52      52      51      53      52       2
## 3              3       2       8      50      50      50      50      50       2
## 4              4       2       1      51      53      54      52      57       2
## 5              4       2       2      48     100     100     100     100       1
## 6              4       2       8      50      50      50      50      50       2
## # ... with 4 more variables: asked_deplane <int>, detained <int>,
## #   age_group <int>, gender <int>, and abbreviated variable names
```

```
## #   1: travel_frequency, 2: discrimination_award, 3: religion_representation,
## #   4: checkin_exp, 5: drop_off_exp, 6: security_exp, 7: boarding_exp,
## #   8: baggage_exp, 9: avoid_travel
```

Statistics Summary

```
summary(travel_df)
```

```
##  travel_frequency discrimination_award religion_representation  checkin_exp
##  Min.   :2.000    Min.   :1.000        Min.   : 1.000          Min.   :  0.00
##  1st Qu.:3.000    1st Qu.:2.000        1st Qu.: 8.000          1st Qu.: 50.00
##  Median :3.000    Median :2.000        Median : 8.000          Median : 50.00
##  Mean   :3.307    Mean   :1.996        Mean   : 7.175          Mean   : 52.78
##  3rd Qu.:4.000    3rd Qu.:2.000        3rd Qu.: 8.000          3rd Qu.: 52.00
##  Max.   :5.000    Max.   :3.000        Max.   :10.000          Max.   :100.00
##   drop_off_exp     security_exp     boarding_exp     baggage_exp
##  Min.   :  0.00   Min.   :  0.00   Min.   :  0.00   Min.   :  0.00
##  1st Qu.: 50.00   1st Qu.: 45.00   1st Qu.: 50.00   1st Qu.: 50.00
##  Median : 50.00   Median : 50.00   Median : 50.00   Median : 50.00
##  Mean   : 54.03   Mean   : 49.33   Mean   : 53.86   Mean   : 54.19
##  3rd Qu.: 51.00   3rd Qu.: 52.25   3rd Qu.: 54.00   3rd Qu.: 52.00
##  Max.   :100.00   Max.   :100.00   Max.   :100.00   Max.   :100.00
##   avoid_travel   asked_deplane      detained       age_group
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :2.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :2.000   Median :2.000   Median :2.000   Median :2.000
##  Mean   :1.925   Mean   :1.991   Mean   :1.895   Mean   :3.263
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:4.000
##  Max.   :3.000   Max.   :3.000   Max.   :3.000   Max.   :7.000
##      gender
##  Min.   :1.000
##  1st Qu.:1.000
##  Median :2.000
##  Mean   :1.737
##  3rd Qu.:2.000
##  Max.   :7.000
```

Correlation Matrix

```
cor(travel_df)
```

```
##                         travel_frequency discrimination_award
## travel_frequency              1.00000000          0.050301983
## discrimination_award          0.05030198          1.000000000
## religion_representation      -0.17042104         -0.064206341
## checkin_exp                  -0.18442616         -0.018131229
## drop_off_exp                 -0.21619138         -0.064583064
## security_exp                 -0.15594661         -0.149649395
## boarding_exp                 -0.15361096          0.041805059
## baggage_exp                  -0.13019796         -0.104751905
## avoid_travel                 -0.01789003         -0.091459273
## asked_deplane                -0.12170409         -0.001661907
## detained                      0.08810392          0.067024953
## age_group                     0.02794714         -0.027403867
```

```
## gender                       -0.02648988          0.142564964
##                         religion_representation checkin_exp drop_off_exp
## travel_frequency                    -0.170421041 -0.18442616 -0.216191381
## discrimination_award                -0.064206341 -0.01813123 -0.064583064
## religion_representation              1.000000000  0.15701258  0.029648479
## checkin_exp                          0.157012577  1.00000000  0.593706474
## drop_off_exp                         0.029648479  0.59370647  1.000000000
## security_exp                         0.240648885  0.51572287  0.593494346
## boarding_exp                         0.080960539  0.54655595  0.640945940
## baggage_exp                         -0.003142728  0.38789415  0.578538671
## avoid_travel                         0.257583291  0.18047291  0.037875263
## asked_deplane                        0.129067753  0.09943460 -0.151613384
## detained                             0.094712975  0.13587279  0.000611184
## age_group                           -0.152158476 -0.02388588  0.068774396
## gender                               0.033963825 -0.07966214 -0.085962703
##                         security_exp boarding_exp  baggage_exp avoid_travel
## travel_frequency         -0.15594661  -0.15361096 -0.130197955  -0.01789003
## discrimination_award     -0.14964939   0.04180506 -0.104751905  -0.09145927
## religion_representation   0.24064888   0.08096054 -0.003142728   0.25758329
## checkin_exp               0.51572287   0.54655595  0.387894150   0.18047291
## drop_off_exp              0.59349435   0.64094594  0.578538671   0.03787526
## security_exp              1.00000000   0.58638119  0.444420307   0.14098646
## boarding_exp              0.58638119   1.00000000  0.609881104  -0.04280326
## baggage_exp               0.44442031   0.60988110  1.000000000  -0.10566007
## avoid_travel              0.14098646  -0.04280326 -0.105660072   1.00000000
## asked_deplane            -0.05910272  -0.07348152 -0.153426105   0.43440378
## detained                  0.09318665   0.04874314 -0.071548244   0.36621366
## age_group                -0.07493566   0.08203833  0.169599329  -0.18344005
## gender                   -0.03174294  -0.04819540 -0.052490249   0.05738760
##                         asked_deplane     detained   age_group      gender
## travel_frequency         -0.121704085  0.088103917  0.02794714 -0.02648988
## discrimination_award     -0.001661907  0.067024953 -0.02740387  0.14256496
## religion_representation   0.129067753  0.094712975 -0.15215848  0.03396383
## checkin_exp               0.099434598  0.135872794 -0.02388588 -0.07966214
## drop_off_exp             -0.151613384  0.000611184  0.06877440 -0.08596270
## security_exp             -0.059102722  0.093186647 -0.07493566 -0.03174294
## boarding_exp             -0.073481518  0.048743138  0.08203833 -0.04819540
## baggage_exp              -0.153426105 -0.071548244  0.16959933 -0.05249025
## avoid_travel              0.434403779  0.366213664 -0.18344005  0.05738760
## asked_deplane             1.000000000  0.177665841 -0.05172889  0.07451606
## detained                  0.177665841  1.000000000 -0.20409755  0.05366507
## age_group                -0.051728893 -0.204097548  1.00000000 -0.11458357
## gender                    0.074516065  0.053665067 -0.11458357  1.00000000
```

Kaiser-Meyer-Oklin Test (KMO)

```
KMO(travel_df)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = travel_df)
## Overall MSA =  0.76
## MSA for each item =
##        travel_frequency    discrimination_award religion_representation
##                    0.67                    0.40                    0.65
```

```
##             checkin_exp            drop_off_exp            security_exp
##                    0.85                    0.80                    0.83
##            boarding_exp             baggage_exp            avoid_travel
##                    0.81                    0.84                    0.59
##           asked_deplane                detained               age_group
##                    0.55                    0.66                    0.66
##                  gender
##                    0.59
```

KMO overall should be .60 or higher to proceed with factor analysis. Thus, we will drop variables that have MSA less than 0.6

```
# create new clean df
travel_df_clean <- travel_df[,-c(2:3, 9:12)]
head(travel_df_clean)
```

```
## # A tibble: 6 x 7
##   travel_frequency checkin_exp drop_off_exp security_exp boardi~1 bagga~2 gender
##              <dbl>       <dbl>        <dbl>        <dbl>    <dbl>   <dbl>  <int>
## 1                4          54           50           71       50      51      1
## 2                4          52           52           51       53      52      2
## 3                3          50           50           50       50      50      2
## 4                4          51           53           54       52      57      1
## 5                4          48          100          100      100     100      2
## 6                4          50           50           50       50      50      2
## # ... with abbreviated variable names 1: boarding_exp, 2: baggage_exp
```

```
# correlation matrix
cor(travel_df_clean)
```

```
##                  travel_frequency checkin_exp drop_off_exp security_exp
## travel_frequency       1.00000000 -0.18442616   -0.2161914  -0.15594661
## checkin_exp           -0.18442616  1.00000000    0.5937065   0.51572287
## drop_off_exp          -0.21619138  0.59370647    1.0000000   0.59349435
## security_exp          -0.15594661  0.51572287    0.5934943   1.00000000
## boarding_exp          -0.15361096  0.54655595    0.6409459   0.58638119
## baggage_exp           -0.13019796  0.38789415    0.5785387   0.44442031
## gender                -0.02648988 -0.07966214   -0.0859627  -0.03174294
##                  boarding_exp baggage_exp      gender
## travel_frequency   -0.1536110 -0.13019796 -0.02648988
## checkin_exp         0.5465560  0.38789415 -0.07966214
## drop_off_exp        0.6409459  0.57853867 -0.08596270
## security_exp        0.5863812  0.44442031 -0.03174294
## boarding_exp        1.0000000  0.60988110 -0.04819540
## baggage_exp         0.6098811  1.00000000 -0.05249025
## gender             -0.0481954 -0.05249025  1.00000000
```

```
KMO(travel_df_clean)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = travel_df_clean)
## Overall MSA =  0.85
## MSA for each item =
## travel_frequency      checkin_exp     drop_off_exp     security_exp
##             0.89             0.87             0.84             0.88
##     boarding_exp      baggage_exp           gender
##             0.84             0.84             0.69
```

```
# Barlett Test
bart_spher(travel_df_clean)
```

```
##  Bartlett's Test of Sphericity
##
## Call: bart_spher(x = travel_df_clean)
##
##      X2 = 505.227
##      df = 21
## p-value < 2.22e-16
```

Principal Component Analysis

```
pca1 = princomp(travel_df_clean, scores=TRUE, cor=TRUE)
summary(pca1)
```

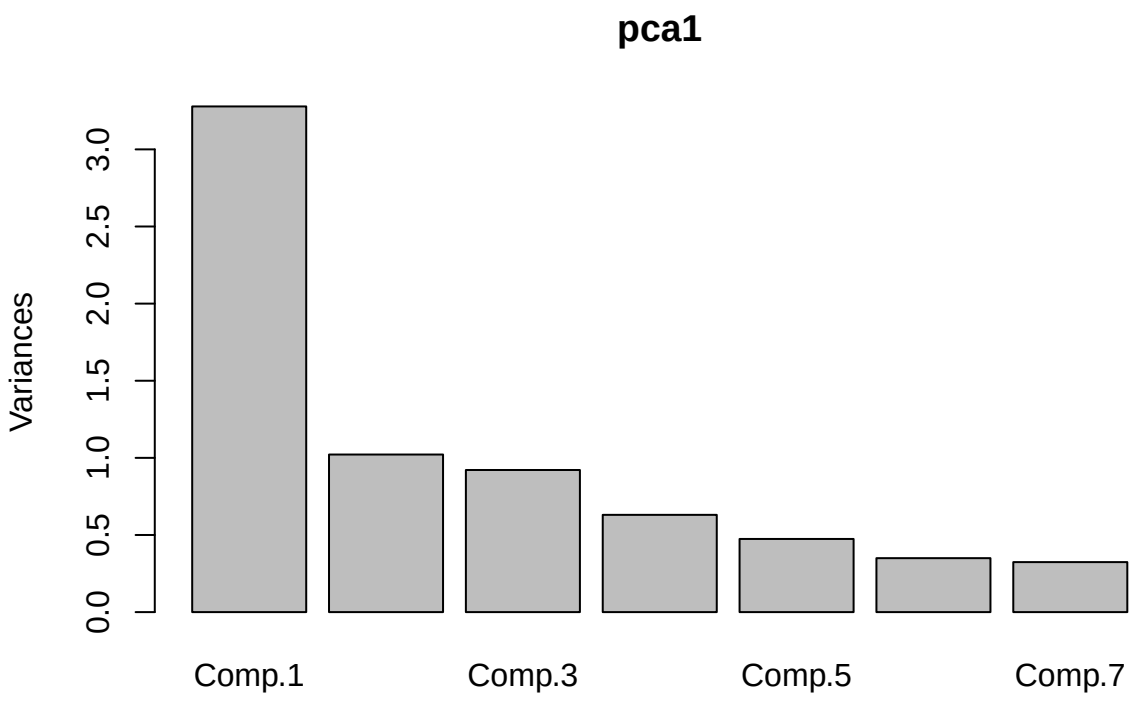
```
## Importance of components:
##                           Comp.1    Comp.2    Comp.3     Comp.4     Comp.5
## Standard deviation     1.8106130 1.0107065 0.9598375 0.79406508 0.68869584
## Proportion of Variance 0.4683314 0.1459325 0.1316126 0.09007705 0.06775742
## Cumulative Proportion  0.4683314 0.6142639 0.7458765 0.83595351 0.90371093
##                           Comp.6     Comp.7
## Standard deviation     0.59143183 0.56941361
## Proportion of Variance 0.04997023 0.04631884
## Cumulative Proportion  0.95368116 1.00000000
```

```
loadings(pca1)
```

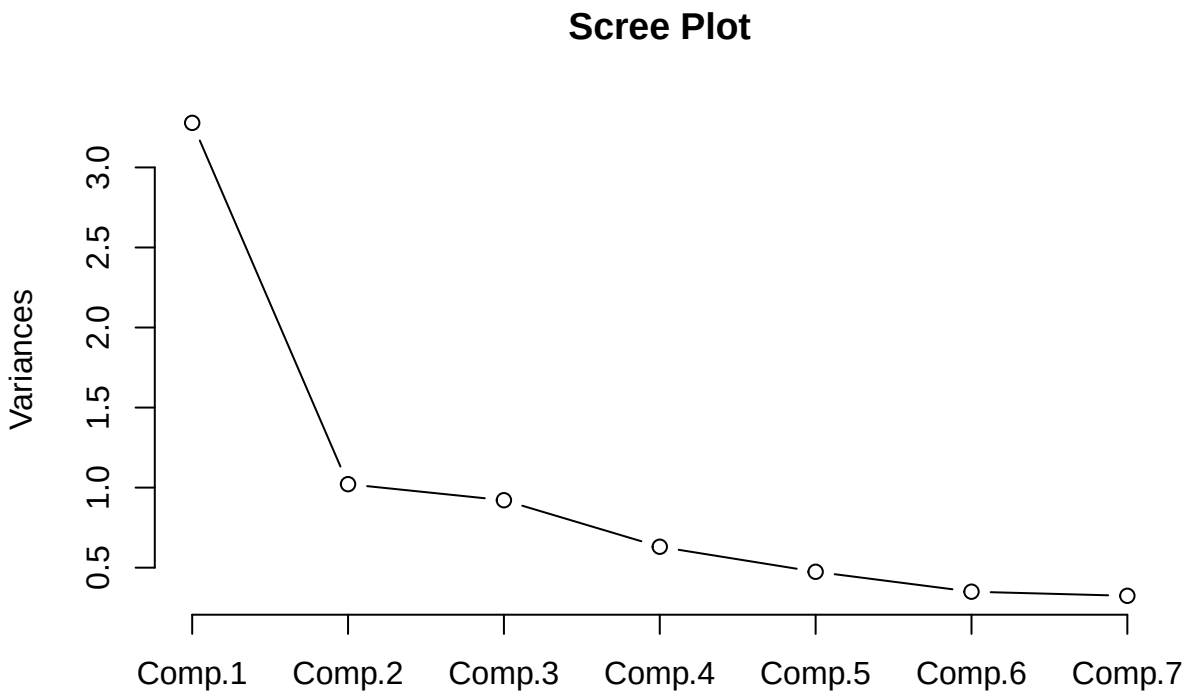
```
##
## Loadings:
##                  Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
## travel_frequency  0.163  0.496  0.840  0.122
## checkin_exp      -0.418                0.592  0.586        -0.353
## drop_off_exp     -0.474                              0.645  0.593
## security_exp     -0.430         0.105  0.312 -0.785        -0.294
## boarding_exp     -0.466         0.143 -0.150        -0.742  0.435
## baggage_exp      -0.411         0.158 -0.716  0.157  0.141 -0.497
## gender                 -0.867  0.487
##
##                Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
## SS loadings     1.000  1.000  1.000  1.000  1.000  1.000  1.000
## Proportion Var  0.143  0.143  0.143  0.143  0.143  0.143  0.143
## Cumulative Var  0.143  0.286  0.429  0.571  0.714  0.857  1.000
```

Scree Plot

```
plot(pca1)
```

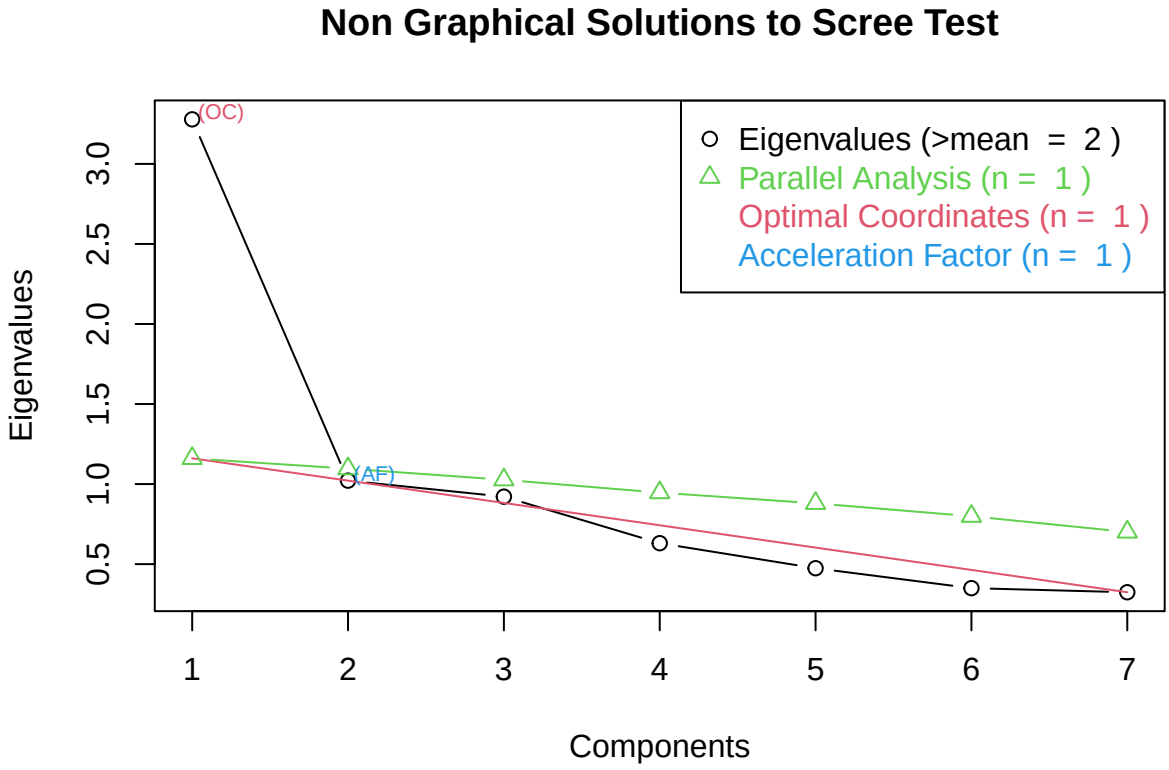



```
screeplot(pca1, type="line", main="Scree Plot")
```



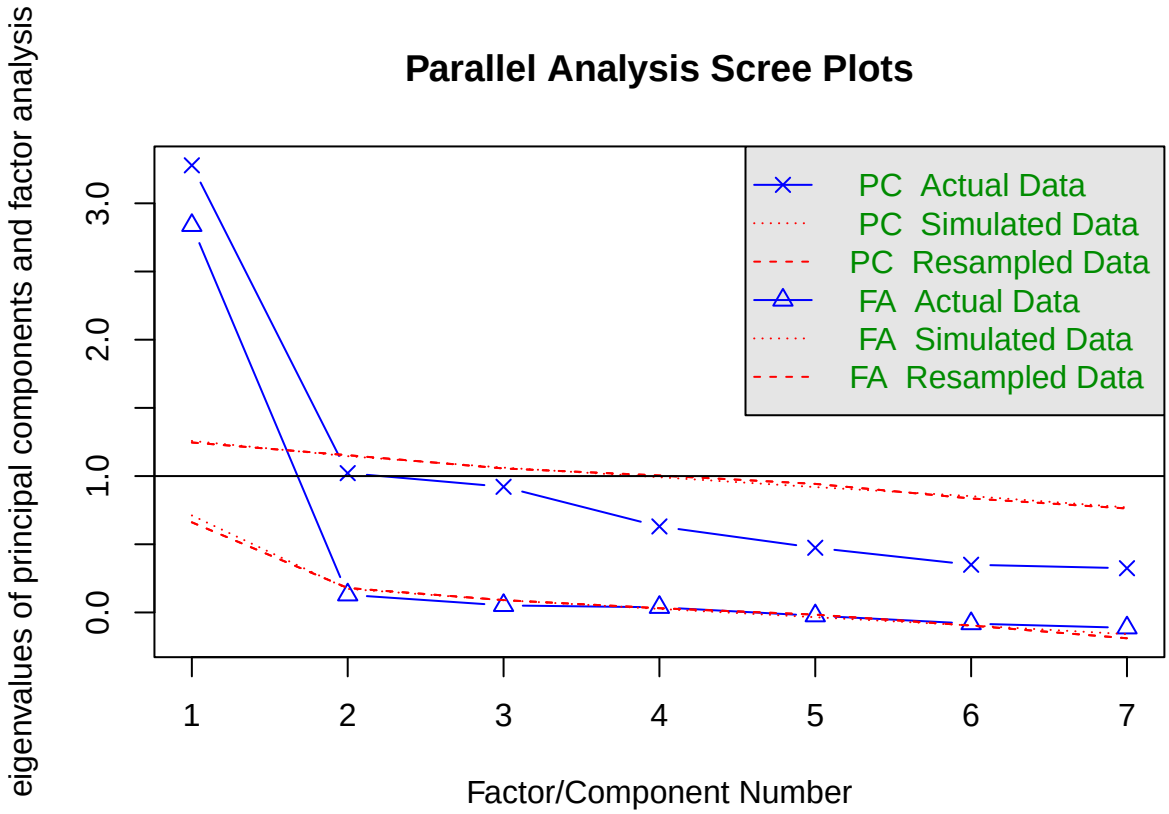
Determine Number of Factors to Retain

```
ev <- eigen(cor(travel_df_clean))
ap <- parallel(subject=nrow(travel_df_clean), var=ncol(travel_df_clean), rep=100,cent=.05)
nS <- nScree(x=ev$values, aparallel=ap$eigen$qevpea)
plotnScree(nS)
```



Exploratory Factorial Analysis

```
fa.parallel(travel_df_clean)
```




```
## Parallel analysis suggests that the number of factors =  1  and the number of components =  1
```

Principal Component

```
fa1 <- principal(travel_df_clean, nfactors=2, rotate="none")
fa1
```

```
## Principal Components Analysis
## Call: principal(r = travel_df_clean, nfactors = 2, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                    PC1   PC2   h2   u2 com
## travel_frequency -0.29 -0.50 0.34 0.66 1.6
## checkin_exp       0.76 -0.02 0.57 0.43 1.0
## drop_off_exp      0.86 -0.02 0.74 0.26 1.0
## security_exp      0.78  0.02 0.61 0.39 1.0
## boarding_exp      0.84 -0.02 0.71 0.29 1.0
## baggage_exp       0.74 -0.04 0.55 0.45 1.0
## gender           -0.10  0.88 0.78 0.22 1.0
##
##                        PC1  PC2
## SS loadings           3.28 1.02
## Proportion Var        0.47 0.15
## Cumulative Var        0.47 0.61
## Proportion Explained  0.76 0.24
## Cumulative Proportion 0.76 1.00
##
## Mean item complexity =  1.1
## Test of the hypothesis that 2 components are sufficient.
```

```
##
## The root mean square of the residuals (RMSR) is  0.11
##  with the empirical chi square  119.46  with prob <  4.3e-22
##
## Fit based upon off diagonal values = 0.92
```

Principal Axis Factor Analysis

```
fa2 <- factor.pa(travel_df_clean, nfactors=2, rotate="none")
```

```
## Warning: factor.pa is deprecated. Please use the fa function with fm=pa
```

```
## maximum iteration exceeded
```

```
fa2
```

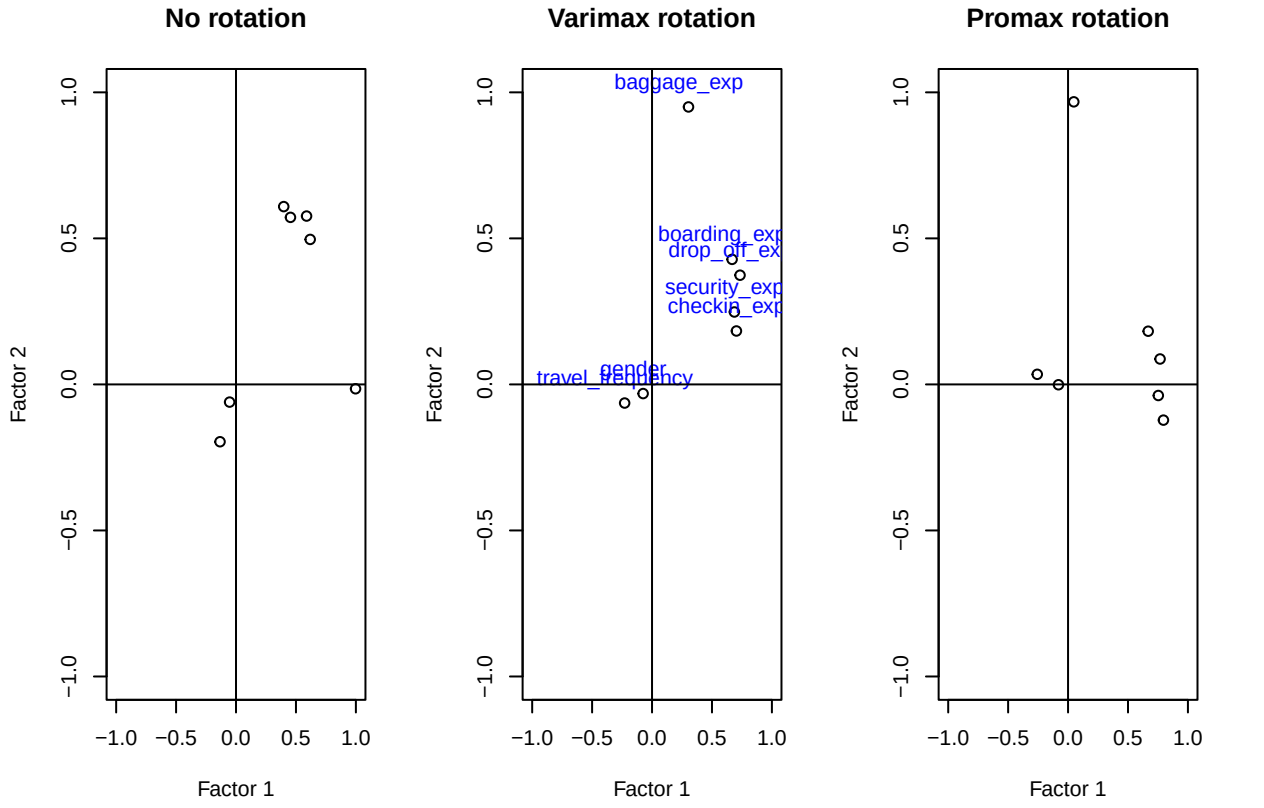
```
## Factor Analysis using method =  pa
## Call: factor.pa(r = travel_df_clean, nfactors = 2, rotate = "none")
## Unstandardized loadings (pattern matrix) based upon covariance matrix
##                     PA1   PA2    h2   u2    H2   U2
## travel_frequency -0.22  0.09 0.057 0.94 0.057 0.94
## checkin_exp       0.69 -0.27 0.553 0.45 0.551 0.45
## drop_off_exp      0.83 -0.07 0.687 0.31 0.689 0.31
## security_exp      0.71 -0.13 0.517 0.48 0.518 0.48
## boarding_exp      0.80  0.06 0.650 0.35 0.650 0.35
## baggage_exp       0.73  0.43 0.713 0.29 0.711 0.29
## gender           -0.08  0.01 0.006 0.99 0.006 0.99
##
##                        PA1  PA2
## SS loadings           2.89 0.29
## Proportion Var        0.41 0.04
## Cumulative Var        0.41 0.46
## Proportion Explained  0.91 0.09
## Cumulative Proportion 0.91 1.00
##
##  Standardized loadings (pattern matrix)
##                   item   PA1   PA2    h2   u2
## travel_frequency    1 -0.22  0.09 0.057 0.94
## checkin_exp         2  0.69 -0.27 0.551 0.45
## drop_off_exp        3  0.83 -0.07 0.689 0.31
## security_exp        4  0.71 -0.13 0.518 0.48
## boarding_exp        5  0.80  0.06 0.650 0.35
## baggage_exp         6  0.73  0.43 0.711 0.29
## gender              7 -0.08  0.01 0.006 0.99
##
##                  PA1  PA2
## SS loadings     2.89 0.29
## Proportion Var  0.41 0.04
## Cumulative Var  0.41 0.45
## Cum. factor Var 0.91 1.00
##
## Mean item complexity =  1.2
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  21  and the objective function was  2.26 with Chi Squa
```

```
## The degrees of freedom for the model are 8  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.02
## The df corrected root mean square of the residuals is  0.03
##
## The harmonic number of observations is  228 with the empirical chi square  2.95  with prob <  0.94
## The total number of observations was  228  with Likelihood Chi Square =  4.47  with prob <  0.81
##
## Tucker Lewis Index of factoring reliability =  1.019
## RMSEA index =  0  and the 90 % confidence intervals are  0 0.049
## BIC =  -38.97
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                  PA1   PA2
## Correlation of (regression) scores with factors   0.94  0.68
## Multiple R square of scores with factors          0.89  0.46
## Minimum correlation of possible factor scores     0.78 -0.08
```

Principal Component Factor Analysis with Rotation

```
fa.none <- factanal(travel_df_clean, factors = 2, rotation = "none")
fa.varimax <- factanal(travel_df_clean, factors = 2, rotation = "varimax")
fa.promax <- factanal(travel_df_clean, factors = 2, rotation = "promax")

par(mfrow = c(1,3))
plot(fa.none$loadings[,1],
     fa.none$loadings[,2],
     xlab = "Factor 1",
     ylab = "Factor 2",
     ylim = c(-1,1),
     xlim = c(-1,1),
     main = "No rotation")
abline(h = 0, v = 0)

plot(fa.varimax$loadings[,1],
     fa.varimax$loadings[,2],
     xlab = "Factor 1",
     ylab = "Factor 2",
     ylim = c(-1,1),
     xlim = c(-1,1),
     main = "Varimax rotation")

text(fa.varimax$loadings[,1]-0.08,
     fa.varimax$loadings[,2]+0.08,
      colnames(travel_df_clean),
      col="blue")
abline(h = 0, v = 0)

plot(fa.promax$loadings[,1],
     fa.promax$loadings[,2],
     xlab = "Factor 1",
     ylab = "Factor 2",
     ylim = c(-1,1),
```

```
     xlim = c(-1,1),
     main = "Promax rotation")
abline(h = 0, v = 0)
```




```
fa.varimax
```

```
##
## Call:
## factanal(x = travel_df_clean, factors = 2, rotation = "varimax")
##
## Uniquenesses:
## travel_frequency      checkin_exp     drop_off_exp     security_exp
##            0.944            0.471            0.321            0.467
##     boarding_exp      baggage_exp           gender
##            0.371            0.005            0.993
##
## Loadings:
##                  Factor1 Factor2
## travel_frequency -0.229
## checkin_exp       0.704   0.183
## drop_off_exp      0.734   0.374
## security_exp      0.687   0.248
## boarding_exp      0.668   0.428
## baggage_exp       0.304   0.950
## gender
##
##                Factor1 Factor2
## SS loadings      2.103   1.326
```

```
## Proportion Var   0.300   0.189
## Cumulative Var   0.300   0.490
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 3.59 on 8 degrees of freedom.
## The p-value is 0.892
```

Exploratory Factor Analysis Using Minimum Residual

```r
faPC_min_res <- fa(r = travel_df_clean, nfactors = 2,rotate = "none")
print(faPC_min_res)
```

```
## Factor Analysis using method =  minres
## Call: fa(r = travel_df_clean, nfactors = 2, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     MR1   MR2     h2     u2 com
## travel_frequency -0.22  0.08 0.0556 0.9444 1.3
## checkin_exp       0.68 -0.26 0.5300 0.4700 1.3
## drop_off_exp      0.82 -0.14 0.6863 0.3137 1.1
## security_exp      0.70 -0.19 0.5256 0.4744 1.1
## boarding_exp      0.79 -0.04 0.6285 0.3715 1.0
## baggage_exp       0.80  0.59 0.9953 0.0047 1.8
## gender           -0.08  0.02 0.0059 0.9941 1.1
##
##                        MR1  MR2
## SS loadings           2.94 0.48
## Proportion Var        0.42 0.07
## Cumulative Var        0.42 0.49
## Proportion Explained  0.86 0.14
## Cumulative Proportion 0.86 1.00
##
## Mean item complexity =  1.2
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  21  and the objective function was  2.26 with Chi Squa
## The degrees of freedom for the model are 8  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.02
## The df corrected root mean square of the residuals is  0.03
##
## The harmonic number of observations is  228 with the empirical chi square  2.85  with prob <  0.94
## The total number of observations was  228  with Likelihood Chi Square =  3.71  with prob <  0.88
##
## Tucker Lewis Index of factoring reliability =  1.023
## RMSEA index =  0  and the 90 % confidence intervals are  0 0.038
## BIC =  -39.73
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                MR1  MR2
## Correlation of (regression) scores with factors   0.96 0.92
## Multiple R square of scores with factors          0.92 0.85
## Minimum correlation of possible factor scores     0.84 0.69
```

```r
faPC_fa <- fa(r = travel_df_clean, nfactors = 2, fm = "pa",rotate = "none")
```

```
## maximum iteration exceeded
```

```r
print(faPC_fa)
```

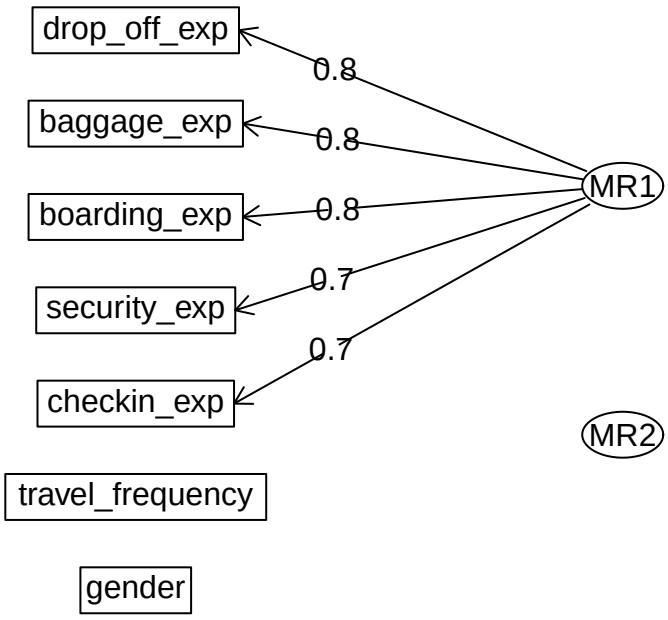
```
## Factor Analysis using method =  pa
## Call: fa(r = travel_df_clean, nfactors = 2, rotate = "none", fm = "pa")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     PA1   PA2    h2   u2 com
## travel_frequency -0.22  0.09 0.057 0.94 1.3
## checkin_exp       0.69 -0.27 0.553 0.45 1.3
## drop_off_exp      0.83 -0.07 0.687 0.31 1.0
## security_exp      0.71 -0.13 0.517 0.48 1.1
## boarding_exp      0.80  0.06 0.650 0.35 1.0
## baggage_exp       0.73  0.43 0.713 0.29 1.6
## gender           -0.08  0.01 0.006 0.99 1.1
##
##                        PA1  PA2
## SS loadings           2.89 0.29
## Proportion Var        0.41 0.04
## Cumulative Var        0.41 0.45
## Proportion Explained  0.91 0.09
## Cumulative Proportion 0.91 1.00
##
## Mean item complexity =  1.2
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  21  and the objective function was  2.26 with Chi Squa
## The degrees of freedom for the model are 8  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.02
## The df corrected root mean square of the residuals is  0.03
##
## The harmonic number of observations is  228 with the empirical chi square  2.95  with prob <  0.94
## The total number of observations was  228  with Likelihood Chi Square =  4.47  with prob <  0.81
##
## Tucker Lewis Index of factoring reliability =  1.019
## RMSEA index =  0  and the 90 % confidence intervals are  0 0.049
## BIC =  -38.97
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                 PA1   PA2
## Correlation of (regression) scores with factors   0.94  0.68
## Multiple R square of scores with factors          0.89  0.46
## Minimum correlation of possible factor scores     0.78 -0.08
```

```r
faPC_ml <- fa(r = travel_df_clean, nfactors = 2, fm="ml", rotate = "none")
print(faPC_ml)
```

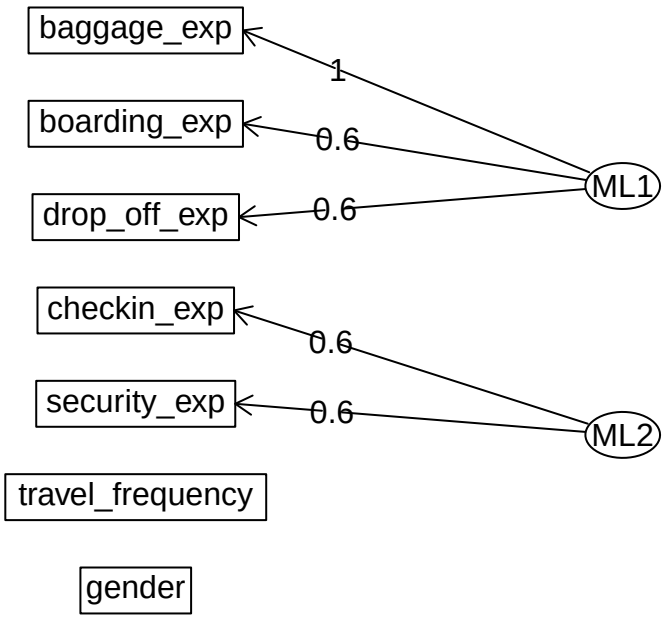
```
## Factor Analysis using method =  ml
## Call: fa(r = travel_df_clean, nfactors = 2, rotate = "none", fm = "ml")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     ML1   ML2     h2    u2 com
## travel_frequency -0.13 -0.20 0.0564 0.944 1.8
```

```
## checkin_exp       0.40  0.61 0.5291 0.471 1.7
## drop_off_exp      0.59  0.58 0.6788 0.321 2.0
## security_exp      0.45  0.57 0.5335 0.467 1.9
## boarding_exp      0.62  0.50 0.6295 0.371 1.9
## baggage_exp       1.00 -0.01 0.9950 0.005 1.0
## gender           -0.05 -0.06 0.0065 0.993 2.0
##
##                        ML1  ML2
## SS loadings           2.11 1.32
## Proportion Var        0.30 0.19
## Cumulative Var        0.30 0.49
## Proportion Explained  0.62 0.38
## Cumulative Proportion 0.62 1.00
##
## Mean item complexity =  1.8
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  21  and the objective function was  2.26 with Chi Squa
## The degrees of freedom for the model are 8  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.02
## The df corrected root mean square of the residuals is  0.03
##
## The harmonic number of observations is  228 with the empirical chi square  2.93  with prob <  0.94
## The total number of observations was  228  with Likelihood Chi Square =  3.59  with prob <  0.89
##
## Tucker Lewis Index of factoring reliability =  1.024
## RMSEA index =  0  and the 90 % confidence intervals are  0 0.035
## BIC =  -39.84
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                 ML1  ML2
## Correlation of (regression) scores with factors   1.00 0.88
## Multiple R square of scores with factors          1.00 0.77
## Minimum correlation of possible factor scores     0.99 0.53
```

```
fa.diagram(faPC_min_res)
```

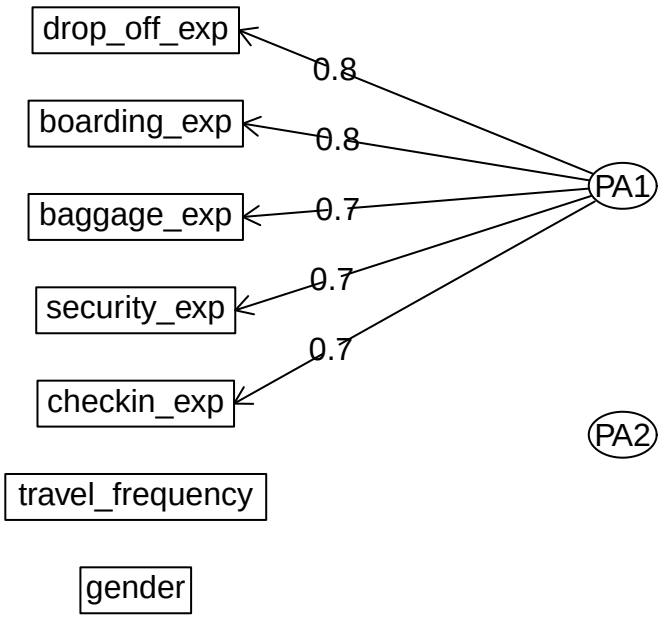



```
fa.diagram(faPC_ml)
```




```
fa.diagram(faPC_fa)
```

Factor Analysis

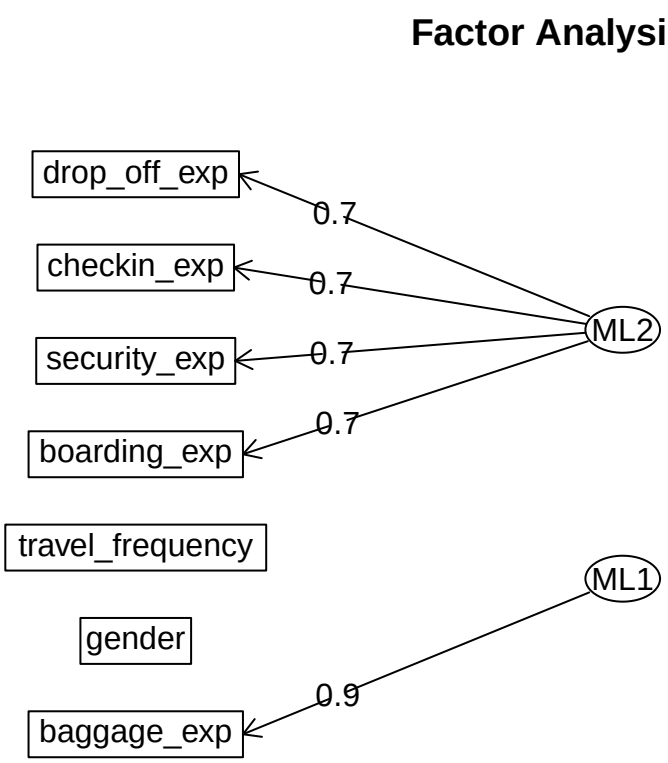



```
faPC_ml_r <- fa(r = travel_df_clean, nfactors = 2, fm="ml", rotate = "varimax")
print(faPC_ml_r)
```

```
## Factor Analysis using method =  ml
## Call: fa(r = travel_df_clean, nfactors = 2, rotate = "varimax", fm = "ml")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     ML2   ML1     h2    u2 com
## travel_frequency -0.23 -0.06 0.0564 0.944 1.2
## checkin_exp       0.70  0.18 0.5291 0.471 1.1
## drop_off_exp      0.73  0.37 0.6788 0.321 1.5
## security_exp      0.69  0.25 0.5335 0.467 1.3
## boarding_exp      0.67  0.43 0.6295 0.371 1.7
## baggage_exp       0.30  0.95 0.9950 0.005 1.2
## gender           -0.07 -0.03 0.0065 0.993 1.3
##
##                        ML2  ML1
## SS loadings           2.10 1.33
## Proportion Var        0.30 0.19
## Cumulative Var        0.30 0.49
## Proportion Explained  0.61 0.39
## Cumulative Proportion 0.61 1.00
##
## Mean item complexity =  1.3
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  21  and the objective function was  2.26 with Chi Squa
## The degrees of freedom for the model are 8  and the objective function was  0.02
##
## The root mean square of the residuals (RMSR) is  0.02
## The df corrected root mean square of the residuals is  0.03
```

```
##
## The harmonic number of observations is  228 with the empirical chi square  2.93  with prob <  0.94
## The total number of observations was  228  with Likelihood Chi Square =  3.59  with prob <  0.89
##
## Tucker Lewis Index of factoring reliability =  1.024
## RMSEA index =  0  and the 90 % confidence intervals are  0 0.035
## BIC =  -39.84
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                  ML2  ML1
## Correlation of (regression) scores with factors   0.89 0.99
## Multiple R square of scores with factors          0.79 0.97
## Minimum correlation of possible factor scores     0.58 0.94
```

```
fa.diagram(faPC_ml_r)
```




```
root.cor.pa <- principal(travel_df_clean, nfactors = 2, rotate = 'varimax')
root.cor.pa
```

```
## Principal Components Analysis
## Call: principal(r = travel_df_clean, nfactors = 2, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     RC1   RC2   h2   u2 com
## travel_frequency -0.25 -0.52 0.34 0.66 1.4
## checkin_exp       0.76  0.04 0.57 0.43 1.0
## drop_off_exp      0.86  0.05 0.74 0.26 1.0
## security_exp      0.77  0.09 0.61 0.39 1.0
## boarding_exp      0.84  0.06 0.71 0.29 1.0
## baggage_exp       0.74  0.03 0.55 0.45 1.0
## gender           -0.18  0.86 0.78 0.22 1.1
##
```

```
##                         RC1  RC2
## SS loadings            3.26 1.04
## Proportion Var         0.47 0.15
## Cumulative Var         0.47 0.61
## Proportion Explained   0.76 0.24
## Cumulative Proportion  0.76 1.00
##
## Mean item complexity =  1.1
## Test of the hypothesis that 2 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.11
##  with the empirical chi square  119.46  with prob <  4.3e-22
##
## Fit based upon off diagonal values = 0.92
```

```
biplot(root.cor.pa)
```