SAGE

Gregorio Ponti

2023-08-23

Intro

Import Libraries

```
library(tidyverse)
library(ggpubr)
library(sjPlot)
theme_set(theme_pubr())

#for statistics
library(car)
library(lme4)
library(lmerTest)

# for EFA
library(psych) #Main FA work
library(corrplot)
library(nFactors) #Help with number of factors to extract
library(FactoMineR) #Additional functions
library(lavaan) #For CFA?
library(sem) #Structural Equation Modeling (used for CFA)
```

Summary

Import -> Tidy Data -> Transform into what we want

EFA process

1. Calculate the Kaiser-Meyer-Olkin (KMO) values for every item. If any items have a KMO below the cutoff value, then the item with the lowest value is removed and the step is repeated. KMO values above 0.6 are kept, though above 0.8 are preferred.
2. KMO measures the suitability for factor analysis by estimating the proportion of variance among all observed variables
3. Check whether the items can be factored using Bartlett's test of sphericity. A low p-score indicates that factor analysis can be performed.
4. Compares the correlation matrix to the identity matrix (checks whether there are correlations)
5. Calculate the EFA model using factoring and a specified number of factors.
6. Calculate the commonalities, which are the proportion of the item's variance explained by the factors. If any item is below the cutoff (<0.2), then the item with the lowest value is dropped and then restart at Step 1.
7. Calculate the item loadings. If there are items that fail to load to any factor, then remove the item with the smallest max loading and then restart at Step 1.

8. Create a model for the CFA by placing each item onto the factor that contains the item's largest loading. If any items load equally onto more than one factor, then add to all factors where this is the case.
9. Fit this model using Confirmatory Factor Analysis to the original data and extract a fit statistic (Akaike information criterion, or similar) to be used as a comparison for the ideal number of factors.
10. Change the number of factors and repeat the above steps.
11. Plot the fit statistic vs the number of factors. The model with the local minimum index is the preferred model.

Data Prep

Import Data

```
df <- read.csv(file = "./ExportedFiles/R_data.csv")
head(df)
```

```
##   X      X0    X1   X2   X3   X4   X5  X6   X7   X8   X9  X10 X11  X12  X13
## 1 0 PHY105N 57135 -1.0  0.5 -0.5 -0.5 0.0  1.0  0.5 -0.5 -1.0 0.0  0.5 -0.5
## 2 1 PHY105N 57135  1.0  0.0 -0.5  0.5 0.5 -1.0  0.5  0.5  1.0 1.0  0.0  0.0
## 3 2 PHY105N 57135  0.5  0.5  0.5  0.5 0.5 -1.0  0.5  0.0  0.5 0.5 -0.5  0.0
## 4 3 PHY105N 57135  0.0  0.0 -1.0  0.0 0.0 -0.5  0.0  0.5  0.5 0.5  0.0  0.0
## 5 4 PHY105N 57135  0.0 -0.5  0.0  0.5 1.0 -1.0  1.0  1.0  1.0 1.0  0.5  0.5
## 6 5 PHY105N 57135  0.0 -0.5 -0.5  0.5 1.0 -1.0 -0.5  0.5  1.0 1.0  0.0  0.0
##    X14 X15 X16  X17  X18  X19  X20  X21 X22 X23  X24 X25  X26  X27  X28 X29 X30
## 1  1.0 0.5 0.5  0.5  1.0  0.5 -1.0 -1.0 0.0 1.0 -1.0 0.0 -1.0 -1.0 -1.0 0.0 0.0
## 2 -0.5 0.0 0.5 -1.0 -1.0 -1.0 -0.5  0.5 1.0 1.0  1.0 0.5 -0.5  1.0 -1.0 1.0 1.0
## 3  0.5 0.5 1.0  0.0 -1.0 -0.5  0.0  1.0 0.5 0.0  0.5 1.0 -0.5  0.5 -0.5 0.5 0.5
## 4 -0.5 0.0 0.5 -0.5 -0.5  0.0  0.0  0.5 1.0 0.5  0.5 0.0  0.0  0.5 -0.5 0.0 0.0
## 5  1.0 1.0 0.5 -1.0 -0.5 -1.0 -1.0  1.0 1.0 0.5  0.0 0.5 -1.0  1.0 -0.5 0.5 0.0
## 6 -1.0 0.5 1.0 -1.0 -1.0 -1.0 -1.0  0.5 0.5 1.0  1.0 1.0 -1.0  1.0 -1.0 0.0 0.5
##   X31  X32 X33    X34 X35                          X36 X37 X38    X39 X40
## 1 1.0 -1.0 1.0   Male                           White
## 2 1.0  1.0 1.0 Female        Black or African American
## 3 1.0  1.0 1.0   Male                           White
## 4 1.0  1.0 1.0   Male                     Asian,White         Korean
## 5 0.6 -0.6 0.6   Male     Hispanic, Latino, or Spanish
## 6 0.6  1.0 0.6 Female     Hispanic, Latino, or Spanish
##                X41 X42
## 1
## 2 African American
## 3
## 4
## 5
## 6
##
## 1
## 2
## 3
## 4
## 5
## 6 Mexican or Mexican American,Some other Hispanic, Latino, or Spanish race or ethnicity\nPrint, for
##        X44 X45 X46 X47 X48                                  X49 X50
## 1                        German,Irish,English,Italian,French
## 2
## 3                                             German,Irish
```

```
## 4                                                    German
## 5
## 6 Spaniard
##                                X51 X52 X53                X54
## 1        Master's degree or above       1 Partner Agreements
## 2               Bachelor's degree       1 Partner Agreements
## 3               Bachelor's degree       1 Partner Agreements
## 4        Master's degree or above       1 Partner Agreements
## 5                     High school       1 Partner Agreements
## 6 Associate's or technical degree       1 Partner Agreements
```

```
describe(df)
```

```
##      vars    n     mean     sd  median  trimmed    mad     min   max  range
## X       1 1273   658.61 379.79   660.0   658.70 487.78     0.0  1315 1315.0
## X0*     2 1273     1.40   0.49     1.0     1.37   0.00     1.0     2    1.0
## X1      3 1273 57010.38 118.17 57020.0 57009.20 148.26 56810.0 57230  420.0
## X2      4 1273     0.26   0.48     0.5     0.28   0.74    -1.0     1    2.0
## X3      5 1273     0.01   0.55     0.0     0.01   0.74    -1.0     1    2.0
## X4      6 1273    -0.19   0.58    -0.5    -0.22   0.74    -1.0     1    2.0
## X5      7 1273     0.50   0.47     0.5     0.56   0.74    -1.0     1    2.0
## X6      8 1273     0.62   0.39     0.5     0.67   0.74    -1.0     1    2.0
## X7      9 1273    -0.49   0.52    -0.5    -0.55   0.74    -1.0     1    2.0
## X8     10 1273     0.48   0.47     0.5     0.54   0.74    -1.0     1    2.0
## X9     11 1273     0.45   0.52     0.5     0.52   0.74    -1.0     1    2.0
## X10    12 1273     0.66   0.41     1.0     0.72   0.00    -1.0     1    2.0
## X11    13 1273     0.66   0.39     0.5     0.72   0.74    -1.0     1    2.0
## X12    14 1273     0.09   0.53     0.0     0.09   0.74    -1.0     1    2.0
## X13    15 1273     0.15   0.49     0.0     0.15   0.74    -1.0     1    2.0
## X14    16 1273    -0.09   0.56     0.0    -0.10   0.74    -1.0     1    2.0
## X15    17 1273     0.33   0.47     0.5     0.36   0.74    -1.0     1    2.0
## X16    18 1273     0.54   0.43     0.5     0.59   0.74    -1.0     1    2.0
## X17    19 1273    -0.52   0.49    -0.5    -0.58   0.74    -1.0     1    2.0
## X18    20 1273    -0.58   0.50    -0.5    -0.65   0.74    -1.5     1    2.5
## X19    21 1273    -0.47   0.51    -0.5    -0.54   0.74    -1.0     1    2.0
## X20    22 1273    -0.26   0.54     0.0    -0.28   0.74    -1.0     1    2.0
## X21    23 1273     0.28   0.57     0.5     0.32   0.74    -1.0     1    2.0
## X22    24 1273     0.20   0.61     0.5     0.23   0.74    -1.0     1    2.0
## X23    25 1273     0.53   0.41     0.5     0.57   0.74    -1.0     1    2.0
## X24    26 1273     0.38   0.44     0.5     0.41   0.74    -1.0     1    2.0
## X25    27 1273     0.58   0.41     0.5     0.63   0.74    -1.0     1    2.0
## X26    28 1273    -0.27   0.57    -0.5    -0.29   0.74    -1.0     1    2.0
## X27    29 1273     0.50   0.49     0.5     0.57   0.74    -1.0     1    2.0
## X28    30 1273    -0.48   0.47    -0.5    -0.54   0.74    -1.0     1    2.0
## X29    31 1273     0.34   0.49     0.5     0.38   0.74    -1.0     1    2.0
## X30    32 1273     0.29   0.49     0.5     0.31   0.74    -1.0     1    2.0
## X31    33 1273     0.55   0.43     0.6     0.60   0.59    -1.0     1    2.0
## X32    34 1273     0.49   0.49     0.6     0.56   0.59    -1.0     1    2.0
## X33    35 1273     0.47   0.47     0.6     0.51   0.59    -1.0     1    2.0
## X34*   36 1273     3.41   1.96     2.0     3.17   0.00     1.0    12   11.0
## X35*   37 1273     1.00   0.10     1.0     1.00   0.00     1.0     4    3.0
## X36*   38 1273    17.61   9.35    20.0    17.65  13.34     1.0    31   30.0
## X37*   39 1273     1.00   0.10     1.0     1.00   0.00     1.0     4    3.0
## X38*   40 1273     1.03   0.41     1.0     1.00   0.00     1.0     9    8.0
## X39*   41 1273     3.13   4.19     1.0     2.10   0.00     1.0    14   13.0
```

```
## X40*   42 1273     1.64   2.85     1.0     1.00   0.00     1.0    19   18.0
## X41*   43 1273     1.20   0.97     1.0     1.00   0.00     1.0     8    7.0
## X42*   44 1273     1.02   0.27     1.0     1.00   0.00     1.0     6    5.0
## X43*   45 1273     2.57   3.17     1.0     1.92   0.00     1.0    16   15.0
## X44*   46 1273     1.34   2.36     1.0     1.00   0.00     1.0    25   24.0
## X45*   47 1273     1.09   0.57     1.0     1.00   0.00     1.0     6    5.0
## X46*   48 1273     1.06   0.57     1.0     1.00   0.00     1.0     7    6.0
## X47*   49 1273     1.00   0.04     1.0     1.00   0.00     1.0     2    1.0
## X48*   50 1273     1.00   0.06     1.0     1.00   0.00     1.0     3    2.0
## X49*   51 1273     9.60  16.51     1.0     5.54   0.00     1.0    53   52.0
## X50*   52 1273     2.80   7.41     1.0     1.00   0.00     1.0    48   47.0
## X51*   53 1273     4.04   1.82     5.0     3.92   1.48     1.0     8    7.0
## X52*   54 1273     1.20   1.56     1.0     1.00   0.00     1.0    19   18.0
## X53    55 1273     0.91   0.82     1.0     0.89   1.48     0.0     2    2.0
## X54*   56 1273     2.03   0.78     2.0     2.03   1.48     1.0     3    2.0
##        skew kurtosis    se
## X      0.00    -1.20 10.64
## X0*    0.41    -1.83  0.01
## X1     0.07    -1.06  3.31
## X2    -0.32    -0.29  0.01
## X3    -0.03    -0.64  0.02
## X4     0.37    -0.71  0.02
## X5    -1.00     0.91  0.01
## X6    -1.09     1.83  0.01
## X7     0.71    -0.30  0.01
## X8    -0.93     0.74  0.01
## X9    -0.97     0.51  0.01
## X10   -1.04     0.58  0.01
## X11   -1.11     1.28  0.01
## X12   -0.09    -0.52  0.01
## X13   -0.15    -0.26  0.01
## X14    0.19    -0.76  0.02
## X15   -0.52     0.07  0.01
## X16   -0.90     0.94  0.01
## X17    0.75    -0.11  0.01
## X18    1.11     0.78  0.01
## X19    0.77    -0.06  0.01
## X20    0.12    -0.79  0.02
## X21   -0.59    -0.55  0.02
## X22   -0.32    -0.95  0.02
## X23   -0.84     0.95  0.01
## X24   -0.54     0.39  0.01
## X25   -0.94     1.04  0.01
## X26    0.21    -1.05  0.02
## X27   -0.90     0.32  0.01
## X28    0.69    -0.16  0.01
## X29   -0.44    -0.29  0.01
## X30   -0.32    -0.39  0.01
## X31   -0.91     0.45  0.01
## X32   -1.16     1.04  0.01
## X33   -0.70    -0.19  0.01
## X34*   1.74     4.48  0.05
## X35*  24.43   630.67  0.00
## X36*  -0.12    -1.55  0.26
```

```
## X37* 24.43   630.67  0.00
## X38* 15.07   243.26  0.01
## X39*  1.86     1.72  0.12
## X40*  4.50    18.93  0.08
## X41*  5.70    32.30  0.03
## X42* 14.29   211.92  0.01
## X43*  2.43     6.39  0.09
## X44*  7.50    58.09  0.07
## X45*  6.33    40.20  0.02
## X46*  9.50    90.97  0.02
## X47* 25.14   630.51  0.00
## X48* 28.64   859.41  0.00
## X49*  1.76     1.51  0.46
## X50*  4.37    18.62  0.21
## X51*  0.23    -0.62  0.05
## X52*  8.23    69.99  0.04
## X53   0.17    -1.49  0.02
## X54* -0.05    -1.37  0.02
```

Process Data