Dr.H

Mo’men Mohamed

March 11, 2019

Loading the important package

library(lavaan) ; library(GPArotation) ; library(knitr) ; library(knitLatex) ; library(readxl) ; library(ggplot2) ;library(tidyr) ; library(dplyr) ; library(magrittr) ; library(car) ; library(psych) ; library(kableExtra)

## Warning: package 'lavaan' was built under R version 3.5.2

## Warning: package 'GPArotation' was built under R version 3.5.2

## Warning: package 'knitLatex' was built under R version 3.5.2

## Warning: package 'car' was built under R version 3.5.2

## Warning: package 'psych' was built under R version 3.5.2

## Warning: package 'kableExtra' was built under R version 3.5.2

Reading the data and importing it

big<-read\_excel(path = "./data sets/big5.xlsx")

Looking at the data structure

str(big)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 225 obs. of 33 variables:  
## $ serial: num 1 2 3 4 5 6 7 8 9 10 ...  
## $ Age : num 18 19 18 19 18 22 22 18 23 19 ...  
## $ Sex : num 1 1 1 1 1 1 1 1 1 1 ...  
## $ E1 : num 3 3 3 2 2 4 1 2 3 2 ...  
## $ E2 : num 3 2 4 3 3 4 4 4 3 1 ...  
## $ E3 : num 2 4 3 2 2 4 4 3 3 2 ...  
## $ E4 : num 3 2 4 2 3 4 4 4 3 4 ...  
## $ E6 : num 4 4 1 3 3 4 1 4 3 2 ...  
## $ E7 : num 3 3 4 3 3 4 4 2 2 4 ...  
## $ N1 : num 2 2 4 4 3 2 4 1 2 2 ...  
## $ N2 : num 1 2 1 2 2 3 2 1 2 1 ...  
## $ N3 : num 1 2 2 2 2 2 1 1 2 1 ...  
## $ N5 : num 3 2 1 2 4 3 4 1 3 4 ...  
## $ N8 : num 4 2 1 2 2 3 4 1 2 2 ...  
## $ N9 : num 1 2 2 2 1 2 4 1 2 4 ...  
## $ A1 : num 3 3 4 3 3 4 1 4 3 1 ...  
## $ A2 : num 3 3 2 4 2 4 2 4 2 1 ...  
## $ A3 : num 4 2 3 4 3 3 3 4 3 2 ...  
## $ A4 : num 3 3 3 4 4 4 3 4 3 2 ...  
## $ A5 : num 3 2 3 4 4 4 4 3 4 2 ...  
## $ A6 : num 3 3 3 4 2 4 1 4 2 2 ...  
## $ O2 : num 3 2 4 3 4 3 4 3 4 2 ...  
## $ O3 : num 3 1 3 1 2 2 1 4 3 2 ...  
## $ O4 : num 1 2 4 3 3 4 4 4 2 2 ...  
## $ O5 : num 2 3 4 3 4 3 1 4 2 2 ...  
## $ O6 : num 3 4 3 3 3 3 4 3 2 2 ...  
## $ O8 : num 4 3 3 3 4 4 4 4 2 2 ...  
## $ C3 : num 4 2 4 3 2 3 4 4 3 2 ...  
## $ C4 : num 2 4 4 3 4 4 4 4 3 2 ...  
## $ C5 : num 3 3 2 3 4 4 1 4 3 2 ...  
## $ C6 : num 4 2 3 2 3 4 1 4 2 2 ...  
## $ C7 : num 2 2 3 3 4 4 1 4 3 3 ...  
## $ C8 : num 2 3 2 3 3 4 1 4 4 1 ...

We need to remove “serial” and convert “sex” to factor Male and female

big<-big[,-1]  
  
big$Sex<-factor(x = big$Sex,levels = c(1,2),labels = c("male","female"))  
  
kable(head(big),format = "markdown")

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | E1 | E2 | E3 | E4 | E6 | E7 | N1 | N2 | N3 | N5 | N8 | N9 | A1 | A2 | A3 | A4 | A5 | A6 | O2 | O3 | O4 | O5 | O6 | O8 | C3 | C4 | C5 | C6 | C7 | C8 |
| 18 | male | 3 | 3 | 2 | 3 | 4 | 3 | 2 | 1 | 1 | 3 | 4 | 1 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 2 | 3 | 4 | 4 | 2 | 3 | 4 | 2 | 2 |
| 19 | male | 3 | 2 | 4 | 2 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | 2 | 3 | 2 | 1 | 2 | 3 | 4 | 3 | 2 | 4 | 3 | 2 | 2 | 3 |
| 18 | male | 3 | 4 | 3 | 4 | 1 | 4 | 4 | 1 | 2 | 1 | 1 | 2 | 4 | 2 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 4 | 3 | 3 | 4 | 4 | 2 | 3 | 3 | 2 |
| 19 | male | 2 | 3 | 2 | 2 | 3 | 3 | 4 | 2 | 2 | 2 | 2 | 2 | 3 | 4 | 4 | 4 | 4 | 4 | 3 | 1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 18 | male | 2 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 2 | 4 | 2 | 1 | 3 | 2 | 3 | 4 | 4 | 2 | 4 | 2 | 3 | 4 | 3 | 4 | 2 | 4 | 4 | 3 | 4 | 3 |
| 22 | male | 4 | 4 | 4 | 4 | 4 | 4 | 2 | 3 | 2 | 3 | 3 | 2 | 4 | 4 | 3 | 4 | 4 | 4 | 3 | 2 | 4 | 3 | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 4 |

Now lets look at some summary for the data over all

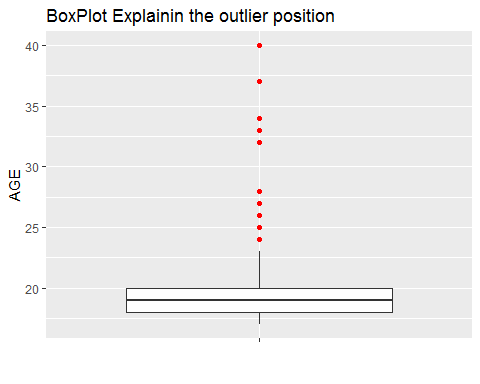
summary(big) #Foor accurecy, out bound items and missing data

## Age Sex E1 E2   
## Min. :17.00 male :114 Min. :1.000 Min. :1.000   
## 1st Qu.:18.00 female:111 1st Qu.:2.000 1st Qu.:2.000   
## Median :19.00 Median :2.000 Median :3.000   
## Mean :20.21 Mean :2.551 Mean :2.538   
## 3rd Qu.:20.00 3rd Qu.:3.000 3rd Qu.:3.000   
## Max. :40.00 Max. :4.000 Max. :4.000   
## E3 E4 E6 E7   
## Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:3.000 1st Qu.:2.000   
## Median :2.000 Median :3.000 Median :3.000 Median :3.000   
## Mean :2.569 Mean :2.667 Mean :3.196 Mean :2.889   
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :4.000 Max. :4.000 Max. :4.000 Max. :4.000   
## N1 N2 N3 N5   
## Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000   
## Median :2.000 Median :2.000 Median :2.00 Median :2.000   
## Mean :2.356 Mean :1.942 Mean :1.96 Mean :2.302   
## 3rd Qu.:3.000 3rd Qu.:2.000 3rd Qu.:2.00 3rd Qu.:3.000   
## Max. :4.000 Max. :4.000 Max. :4.00 Max. :4.000   
## N8 N9 A1 A2   
## Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:3.000 1st Qu.:2.000   
## Median :2.000 Median :2.000 Median :4.000 Median :3.000   
## Mean :2.013 Mean :1.973 Mean :3.293 Mean :2.893   
## 3rd Qu.:3.000 3rd Qu.:2.000 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :4.000 Max. :4.000 Max. :4.000 Max. :4.000   
## A3 A4 A5 A6   
## Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:3.000 1st Qu.:3.000 1st Qu.:2.000   
## Median :3.000 Median :4.000 Median :3.000 Median :3.000   
## Mean :3.142 Mean :3.493 Mean :3.342 Mean :3.138   
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :4.000 Max. :4.000 Max. :4.000 Max. :4.000   
## O2 O3 O4 O5   
## Min. :1.000 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000   
## Median :3.000 Median :3.000 Median :3.000 Median :3.000   
## Mean :2.769 Mean :2.533 Mean :2.693 Mean :3.062   
## 3rd Qu.:4.000 3rd Qu.:3.000 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :4.000 Max. :4.000 Max. :4.000 Max. :4.000   
## O6 O8 C3 C4   
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:2.00 1st Qu.:3.000 1st Qu.:2.000 1st Qu.:2.000   
## Median :3.00 Median :4.000 Median :3.000 Median :3.000   
## Mean :2.88 Mean :3.338 Mean :2.782 Mean :3.071   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :4.00 Max. :4.000 Max. :4.000 Max. :4.000   
## C5 C6 C7 C8   
## Min. :1.00 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:2.00 1st Qu.:2.000 1st Qu.:2.00 1st Qu.:3.000   
## Median :3.00 Median :3.000 Median :3.00 Median :3.000   
## Mean :3.04 Mean :2.844 Mean :2.84 Mean :2.973   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:4.00 3rd Qu.:4.000   
## Max. :4.00 Max. :4.000 Max. :4.00 Max. :4.000

outidx<-which(big$Age>qnorm(p = .99,mean = mean(big$Age),sd = sd(big$Age))) # testing outlier over Propapility 99%  
  
kable(big[outidx,],format = "markdown")

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | E1 | E2 | E3 | E4 | E6 | E7 | N1 | N2 | N3 | N5 | N8 | N9 | A1 | A2 | A3 | A4 | A5 | A6 | O2 | O3 | O4 | O5 | O6 | O8 | C3 | C4 | C5 | C6 | C7 | C8 |
| 40 | male | 2 | 1 | 2 | 3 | 4 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 3 | 3 | 3 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 4 | 4 |
| 40 | male | 3 | 2 | 4 | 4 | 4 | 4 | 2 | 2 | 1 | 2 | 1 | 1 | 4 | 2 | 4 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 3 | 4 |
| 37 | male | 4 | 3 | 4 | 2 | 4 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 3 | 4 | 4 |
| 34 | male | 2 | 1 | 1 | 2 | 3 | 2 | 2 | 1 | 2 | 2 | 1 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 32 | male | 3 | 2 | 3 | 3 | 4 | 3 | 3 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | 3 | 2 | 3 | 3 |
| 32 | female | 3 | 2 | 2 | 2 | 4 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 3 | 2 | 3 | 3 | 3 | 2 | 4 |
| 37 | female | 3 | 1 | 2 | 4 | 4 | 3 | 2 | 3 | 1 | 2 | 1 | 1 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 1 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 33 | female | 2 | 2 | 2 | 2 | 3 | 2 | 3 | 3 | 2 | 2 | 1 | 2 | 4 | 2 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 4 | 4 | 3 | 3 | 4 | 4 | 4 |

big%>%ggplot(aes(x = "",y = Age))+geom\_boxplot(outlier.color = "red")+  
 labs(title = "BoxPlot Explainin the outlier position")+  
 xlab("")+ylab("AGE")



Great , the there is no out of bound items or missing data , and the number of males and females is all most equal

but unfortunately we have some outlier data records in the variable of age

.99 Quantile suggest all age above 28 , and box plot suggest all over 25 years

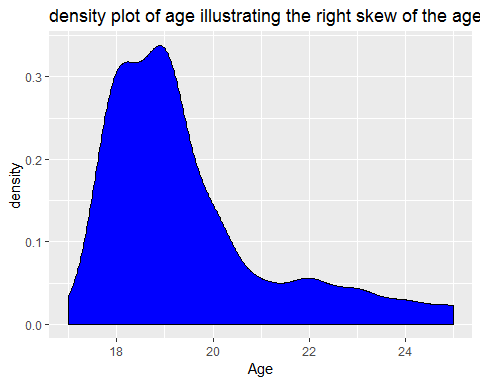
I will remove all about 25 to confirm that all of my sample are in period of early adulthood

I Will remove these outlier

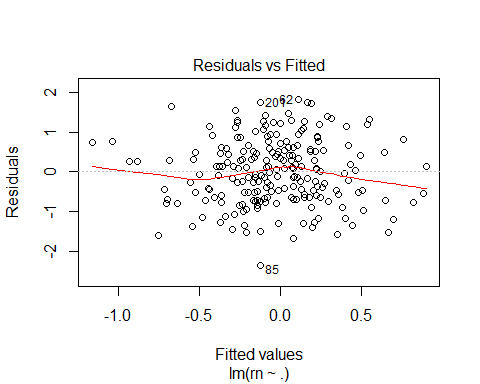
big<-big[!big$Age>25,]

The normality and linearity assumption is critical for Factor analysis and SEM , So, Lets test it

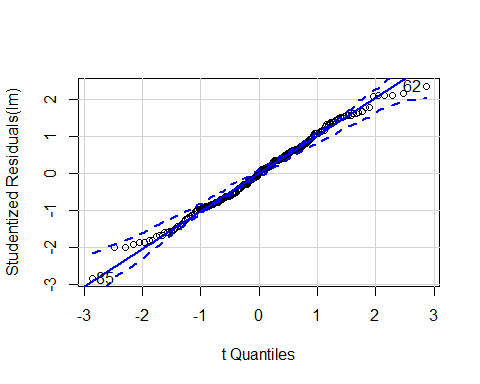
ggplot(data = big,aes(Age))+geom\_density(fill="blue")+labs(title = "density plot of age illustrating the right skew of the age") #Normality inspection of age



set.seed(0124)  
rn<-rnorm(n = nrow(big))  
lm<-lm(rn~.,data = big[,-c(1,2)])  
  
plot(lm,1) #Homogeniety



qqPlot(lm) # Normality and linearity



## [1] 62 85

From The QQplot we can see that we can assume the normality and linearity of the data

Now It is the time to dive more in our data

I need to classify sample age into to grouped

1- Early adults => > 20 y

2- Late Adolescents => < 20 y

big$age\_group<-sapply(big$Age,function(x){  
 if(x>20){paste("Adult")}  
 else{paste("adol")}  
})  
  
kable(head(big),format = "markdown")

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | E1 | E2 | E3 | E4 | E6 | E7 | N1 | N2 | N3 | N5 | N8 | N9 | A1 | A2 | A3 | A4 | A5 | A6 | O2 | O3 | O4 | O5 | O6 | O8 | C3 | C4 | C5 | C6 | C7 | C8 | age\_group |
| 18 | male | 3 | 3 | 2 | 3 | 4 | 3 | 2 | 1 | 1 | 3 | 4 | 1 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 2 | 3 | 4 | 4 | 2 | 3 | 4 | 2 | 2 | adol |
| 19 | male | 3 | 2 | 4 | 2 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | 2 | 3 | 2 | 1 | 2 | 3 | 4 | 3 | 2 | 4 | 3 | 2 | 2 | 3 | adol |
| 18 | male | 3 | 4 | 3 | 4 | 1 | 4 | 4 | 1 | 2 | 1 | 1 | 2 | 4 | 2 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 4 | 3 | 3 | 4 | 4 | 2 | 3 | 3 | 2 | adol |
| 19 | male | 2 | 3 | 2 | 2 | 3 | 3 | 4 | 2 | 2 | 2 | 2 | 2 | 3 | 4 | 4 | 4 | 4 | 4 | 3 | 1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | adol |
| 18 | male | 2 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 2 | 4 | 2 | 1 | 3 | 2 | 3 | 4 | 4 | 2 | 4 | 2 | 3 | 4 | 3 | 4 | 2 | 4 | 4 | 3 | 4 | 3 | adol |
| 22 | male | 4 | 4 | 4 | 4 | 4 | 4 | 2 | 3 | 2 | 3 | 3 | 2 | 4 | 4 | 3 | 4 | 4 | 4 | 3 | 2 | 4 | 3 | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | Adult |

Now lets illustrate the frequency with probabilities of gender in each age class due to decide if we can consider the age as a group for conducting the measurement variance proving procedures

tab<-table(big$Sex,big$age\_group)  
  
kable(round(prop.table(tab,margin = 1),2),format = "markdown",caption = "propapility table of age and sex groups ")

|  |  |  |
| --- | --- | --- |
|  | adol | Adult |
| male | 0.78 | 0.22 |
| female | 0.83 | 0.17 |

As we see , we can’t assume that the age group is a valid group as the sex is not uniformly distributed in it !!

But we still need to confirm that the two sex have the same mean of age

we will prove that using independent t-test

with(big,  
 bartlett.test(Age~Sex)) # the variance is equal between group

##   
## Bartlett test of homogeneity of variances  
##   
## data: Age by Sex  
## Bartlett's K-squared = 3.104, df = 1, p-value = 0.0781

with(big,  
 t.test(formula=Age~Sex,alternative="two.sided",conf.level=.95,paired=F))

##   
## Welch Two Sample t-test  
##   
## data: Age by Sex  
## t = -0.21163, df = 203.01, p-value = 0.8326  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.5349413 0.4312376  
## sample estimates:  
## mean in group male mean in group female   
## 19.48148 19.53333

Congrats!! the two groups has the same mean of age with p.value 0.8326 and equal in variance with P, Value =0.0781

The following steps include conducting EFA and CFA

Lets start with EFA:

Firstly lets test the assumptions : **Additivity** , \*\* sample adequacy \*\*

#Additivity   
  
corr<-cor(x = big[,-c(1,2,ncol(big))],method = "p")  
  
kable(corr,digits = 2,format = "markdown") #correlation table

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | E1 | E2 | E3 | E4 | E6 | E7 | N1 | N2 | N3 | N5 | N8 | N9 | A1 | A2 | A3 | A4 | A5 | A6 | O2 | O3 | O4 | O5 | O6 | O8 | C3 | C4 | C5 | C6 | C7 | C8 |
| E1 | 1.00 | 0.38 | 0.55 | 0.43 | 0.43 | 0.36 | -0.11 | -0.08 | -0.20 | -0.17 | -0.02 | -0.19 | 0.02 | 0.09 | 0.08 | 0.10 | 0.18 | 0.18 | 0.08 | 0.05 | 0.01 | 0.14 | 0.11 | 0.23 | 0.10 | 0.16 | 0.18 | 0.06 | 0.11 | 0.07 |
| E2 | 0.38 | 1.00 | 0.43 | 0.27 | 0.34 | 0.39 | -0.09 | -0.19 | -0.27 | -0.13 | -0.04 | -0.27 | 0.05 | 0.07 | 0.12 | 0.14 | 0.24 | 0.20 | 0.13 | 0.07 | 0.19 | 0.07 | 0.11 | 0.33 | 0.09 | 0.15 | 0.26 | 0.06 | 0.14 | 0.10 |
| E3 | 0.55 | 0.43 | 1.00 | 0.34 | 0.42 | 0.47 | -0.12 | -0.21 | -0.32 | -0.13 | -0.07 | -0.22 | 0.10 | 0.07 | 0.06 | 0.13 | 0.25 | 0.18 | 0.16 | 0.14 | 0.22 | 0.20 | 0.20 | 0.27 | 0.24 | 0.24 | 0.18 | 0.16 | 0.27 | 0.17 |
| E4 | 0.43 | 0.27 | 0.34 | 1.00 | 0.26 | 0.38 | -0.07 | -0.11 | -0.19 | -0.02 | 0.01 | -0.09 | -0.01 | 0.15 | 0.13 | 0.16 | 0.23 | 0.22 | 0.12 | 0.13 | 0.16 | 0.19 | 0.11 | 0.19 | 0.09 | 0.23 | 0.11 | 0.03 | 0.08 | 0.06 |
| E6 | 0.43 | 0.34 | 0.42 | 0.26 | 1.00 | 0.35 | -0.15 | -0.31 | -0.44 | -0.33 | -0.15 | -0.48 | -0.05 | 0.09 | -0.02 | 0.07 | 0.07 | 0.09 | -0.09 | 0.06 | -0.03 | 0.09 | 0.09 | 0.16 | 0.01 | 0.06 | 0.16 | 0.07 | 0.04 | 0.16 |
| E7 | 0.36 | 0.39 | 0.47 | 0.38 | 0.35 | 1.00 | -0.06 | -0.22 | -0.33 | -0.14 | -0.04 | -0.26 | 0.10 | 0.07 | 0.13 | 0.20 | 0.29 | 0.21 | 0.18 | 0.11 | 0.13 | 0.21 | 0.17 | 0.30 | 0.17 | 0.32 | 0.29 | 0.00 | 0.17 | 0.06 |
| N1 | -0.11 | -0.09 | -0.12 | -0.07 | -0.15 | -0.06 | 1.00 | 0.40 | 0.34 | 0.49 | 0.33 | 0.16 | 0.03 | 0.08 | 0.01 | 0.04 | 0.05 | 0.00 | -0.11 | 0.03 | -0.10 | -0.10 | -0.10 | 0.02 | -0.03 | -0.11 | -0.09 | -0.14 | -0.09 | -0.22 |
| N2 | -0.08 | -0.19 | -0.21 | -0.11 | -0.31 | -0.22 | 0.40 | 1.00 | 0.56 | 0.44 | 0.28 | 0.40 | 0.10 | 0.04 | 0.08 | 0.00 | 0.07 | 0.09 | 0.01 | -0.03 | -0.03 | -0.11 | -0.08 | -0.07 | 0.01 | -0.14 | -0.04 | -0.01 | -0.03 | -0.04 |
| N3 | -0.20 | -0.27 | -0.32 | -0.19 | -0.44 | -0.33 | 0.34 | 0.56 | 1.00 | 0.36 | 0.28 | 0.51 | 0.05 | 0.08 | 0.12 | 0.03 | 0.02 | 0.01 | 0.02 | -0.02 | -0.02 | -0.08 | -0.03 | -0.14 | 0.07 | -0.18 | -0.08 | 0.02 | -0.08 | -0.11 |
| N5 | -0.17 | -0.13 | -0.13 | -0.02 | -0.33 | -0.14 | 0.49 | 0.44 | 0.36 | 1.00 | 0.51 | 0.36 | -0.03 | -0.03 | 0.03 | -0.04 | 0.00 | -0.07 | -0.14 | -0.01 | -0.01 | -0.17 | -0.14 | -0.02 | 0.00 | -0.11 | -0.10 | -0.11 | -0.08 | -0.14 |
| N8 | -0.02 | -0.04 | -0.07 | 0.01 | -0.15 | -0.04 | 0.33 | 0.28 | 0.28 | 0.51 | 1.00 | 0.30 | -0.08 | 0.03 | 0.04 | 0.00 | 0.07 | -0.02 | -0.10 | -0.08 | -0.02 | -0.13 | -0.04 | 0.03 | 0.06 | -0.14 | -0.11 | -0.11 | -0.18 | -0.21 |
| N9 | -0.19 | -0.27 | -0.22 | -0.09 | -0.48 | -0.26 | 0.16 | 0.40 | 0.51 | 0.36 | 0.30 | 1.00 | 0.04 | 0.01 | 0.10 | 0.02 | -0.04 | -0.06 | 0.02 | 0.01 | -0.01 | -0.10 | -0.05 | -0.08 | 0.05 | -0.11 | -0.14 | -0.08 | -0.05 | -0.12 |
| A1 | 0.02 | 0.05 | 0.10 | -0.01 | -0.05 | 0.10 | 0.03 | 0.10 | 0.05 | -0.03 | -0.08 | 0.04 | 1.00 | 0.33 | 0.39 | 0.36 | 0.33 | 0.36 | 0.13 | 0.25 | 0.20 | 0.16 | 0.00 | 0.18 | 0.20 | 0.16 | 0.15 | 0.32 | 0.35 | 0.19 |
| A2 | 0.09 | 0.07 | 0.07 | 0.15 | 0.09 | 0.07 | 0.08 | 0.04 | 0.08 | -0.03 | 0.03 | 0.01 | 0.33 | 1.00 | 0.47 | 0.31 | 0.34 | 0.36 | 0.07 | 0.08 | 0.12 | 0.16 | 0.14 | 0.14 | 0.13 | 0.10 | 0.10 | 0.11 | 0.17 | 0.14 |
| A3 | 0.08 | 0.12 | 0.06 | 0.13 | -0.02 | 0.13 | 0.01 | 0.08 | 0.12 | 0.03 | 0.04 | 0.10 | 0.39 | 0.47 | 1.00 | 0.42 | 0.47 | 0.36 | 0.12 | 0.19 | 0.18 | 0.16 | 0.12 | 0.25 | 0.24 | 0.07 | 0.15 | 0.20 | 0.26 | 0.10 |
| A4 | 0.10 | 0.14 | 0.13 | 0.16 | 0.07 | 0.20 | 0.04 | 0.00 | 0.03 | -0.04 | 0.00 | 0.02 | 0.36 | 0.31 | 0.42 | 1.00 | 0.52 | 0.40 | 0.07 | 0.20 | 0.22 | 0.24 | 0.16 | 0.24 | 0.17 | 0.18 | 0.22 | 0.17 | 0.26 | 0.17 |
| A5 | 0.18 | 0.24 | 0.25 | 0.23 | 0.07 | 0.29 | 0.05 | 0.07 | 0.02 | 0.00 | 0.07 | -0.04 | 0.33 | 0.34 | 0.47 | 0.52 | 1.00 | 0.49 | 0.23 | 0.18 | 0.30 | 0.22 | 0.17 | 0.28 | 0.30 | 0.34 | 0.36 | 0.24 | 0.36 | 0.19 |
| A6 | 0.18 | 0.20 | 0.18 | 0.22 | 0.09 | 0.21 | 0.00 | 0.09 | 0.01 | -0.07 | -0.02 | -0.06 | 0.36 | 0.36 | 0.36 | 0.40 | 0.49 | 1.00 | 0.13 | 0.19 | 0.32 | 0.21 | 0.18 | 0.19 | 0.22 | 0.29 | 0.36 | 0.36 | 0.27 | 0.29 |
| O2 | 0.08 | 0.13 | 0.16 | 0.12 | -0.09 | 0.18 | -0.11 | 0.01 | 0.02 | -0.14 | -0.10 | 0.02 | 0.13 | 0.07 | 0.12 | 0.07 | 0.23 | 0.13 | 1.00 | 0.41 | 0.47 | 0.47 | 0.34 | 0.08 | 0.29 | 0.31 | 0.26 | 0.20 | 0.30 | 0.26 |
| O3 | 0.05 | 0.07 | 0.14 | 0.13 | 0.06 | 0.11 | 0.03 | -0.03 | -0.02 | -0.01 | -0.08 | 0.01 | 0.25 | 0.08 | 0.19 | 0.20 | 0.18 | 0.19 | 0.41 | 1.00 | 0.29 | 0.35 | 0.22 | 0.20 | 0.23 | 0.13 | 0.09 | 0.18 | 0.22 | 0.17 |
| O4 | 0.01 | 0.19 | 0.22 | 0.16 | -0.03 | 0.13 | -0.10 | -0.03 | -0.02 | -0.01 | -0.02 | -0.01 | 0.20 | 0.12 | 0.18 | 0.22 | 0.30 | 0.32 | 0.47 | 0.29 | 1.00 | 0.53 | 0.35 | 0.26 | 0.34 | 0.24 | 0.29 | 0.26 | 0.26 | 0.24 |
| O5 | 0.14 | 0.07 | 0.20 | 0.19 | 0.09 | 0.21 | -0.10 | -0.11 | -0.08 | -0.17 | -0.13 | -0.10 | 0.16 | 0.16 | 0.16 | 0.24 | 0.22 | 0.21 | 0.47 | 0.35 | 0.53 | 1.00 | 0.35 | 0.29 | 0.22 | 0.23 | 0.28 | 0.19 | 0.30 | 0.30 |
| O6 | 0.11 | 0.11 | 0.20 | 0.11 | 0.09 | 0.17 | -0.10 | -0.08 | -0.03 | -0.14 | -0.04 | -0.05 | 0.00 | 0.14 | 0.12 | 0.16 | 0.17 | 0.18 | 0.34 | 0.22 | 0.35 | 0.35 | 1.00 | 0.25 | 0.37 | 0.27 | 0.26 | 0.29 | 0.28 | 0.26 |
| O8 | 0.23 | 0.33 | 0.27 | 0.19 | 0.16 | 0.30 | 0.02 | -0.07 | -0.14 | -0.02 | 0.03 | -0.08 | 0.18 | 0.14 | 0.25 | 0.24 | 0.28 | 0.19 | 0.08 | 0.20 | 0.26 | 0.29 | 0.25 | 1.00 | 0.19 | 0.23 | 0.28 | 0.18 | 0.20 | 0.08 |
| C3 | 0.10 | 0.09 | 0.24 | 0.09 | 0.01 | 0.17 | -0.03 | 0.01 | 0.07 | 0.00 | 0.06 | 0.05 | 0.20 | 0.13 | 0.24 | 0.17 | 0.30 | 0.22 | 0.29 | 0.23 | 0.34 | 0.22 | 0.37 | 0.19 | 1.00 | 0.38 | 0.36 | 0.47 | 0.45 | 0.24 |
| C4 | 0.16 | 0.15 | 0.24 | 0.23 | 0.06 | 0.32 | -0.11 | -0.14 | -0.18 | -0.11 | -0.14 | -0.11 | 0.16 | 0.10 | 0.07 | 0.18 | 0.34 | 0.29 | 0.31 | 0.13 | 0.24 | 0.23 | 0.27 | 0.23 | 0.38 | 1.00 | 0.59 | 0.29 | 0.37 | 0.30 |
| C5 | 0.18 | 0.26 | 0.18 | 0.11 | 0.16 | 0.29 | -0.09 | -0.04 | -0.08 | -0.10 | -0.11 | -0.14 | 0.15 | 0.10 | 0.15 | 0.22 | 0.36 | 0.36 | 0.26 | 0.09 | 0.29 | 0.28 | 0.26 | 0.28 | 0.36 | 0.59 | 1.00 | 0.42 | 0.50 | 0.36 |
| C6 | 0.06 | 0.06 | 0.16 | 0.03 | 0.07 | 0.00 | -0.14 | -0.01 | 0.02 | -0.11 | -0.11 | -0.08 | 0.32 | 0.11 | 0.20 | 0.17 | 0.24 | 0.36 | 0.20 | 0.18 | 0.26 | 0.19 | 0.29 | 0.18 | 0.47 | 0.29 | 0.42 | 1.00 | 0.66 | 0.42 |
| C7 | 0.11 | 0.14 | 0.27 | 0.08 | 0.04 | 0.17 | -0.09 | -0.03 | -0.08 | -0.08 | -0.18 | -0.05 | 0.35 | 0.17 | 0.26 | 0.26 | 0.36 | 0.27 | 0.30 | 0.22 | 0.26 | 0.30 | 0.28 | 0.20 | 0.45 | 0.37 | 0.50 | 0.66 | 1.00 | 0.41 |
| C8 | 0.07 | 0.10 | 0.17 | 0.06 | 0.16 | 0.06 | -0.22 | -0.04 | -0.11 | -0.14 | -0.21 | -0.12 | 0.19 | 0.14 | 0.10 | 0.17 | 0.19 | 0.29 | 0.26 | 0.17 | 0.24 | 0.30 | 0.26 | 0.08 | 0.24 | 0.30 | 0.36 | 0.42 | 0.41 | 1.00 |

corr%>%apply(MARGIN = 2,function(x){  
 x<1&x>=abs(.9)  
})%>%any() #Testing if there are a correlation over .9 between the items to avoid multicolinearity

## [1] FALSE

#Enough correlation & Sample adequacy  
cortest.bartlett(R = corr,n = nrow(big))$p.value#correlation bartelett test

## [1] 3.919203e-241

KMO(corr)$MSA #Kaiser-Meyer-Olkin test

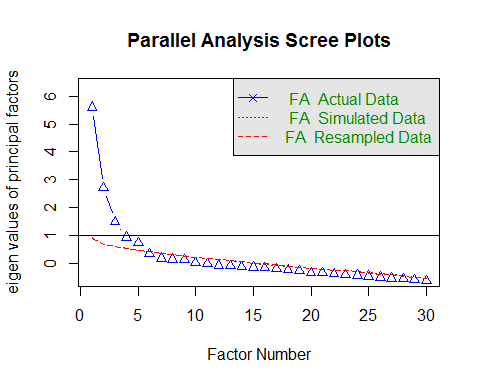
## [1] 0.8202394

since the assumptions have been met ( sig cor bartelett and MSA >.8 and no additivity ), Now we can run EFA

AS there are a priori theory which assume the presence of 5 factors on which the items load in unique way , there is no need to test the number of factors

but lets do it for fun !!

facno<-fa.parallel(x = big[,-c(1,2,ncol(big))],fm = "ml",fa = "fa") # as th data is normal , we choosed the Muximum likelihood method of Extraction



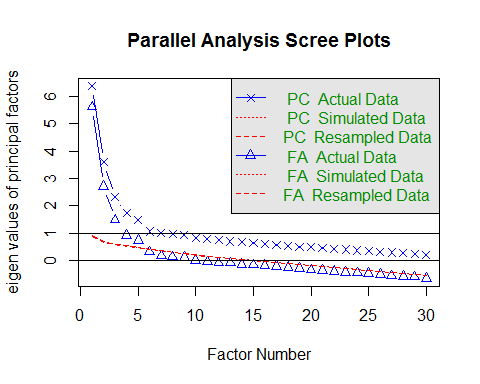
## Parallel analysis suggests that the number of factors = 5 and the number of components = NA

The parallel test suggest 5 factors !! agreeing with the theory , but for fancy lets look at scree plot & Kaiser Criterion

(facno$fa.values>.7)%>%sum() #Kaiser criterion suggest 5 Factors !!

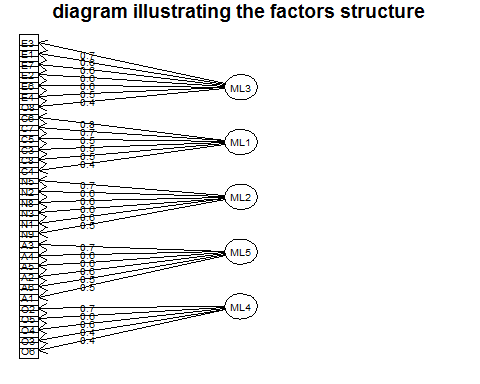
## [1] 5

plot(facno) #Scree plot Suggest 4 Factors with almost more one !!



As we see that Kaiser criterion of Eigen value suggest 5 factors and scree suggests 4 and almost one !! so we will consider those 5 !!

fit<-fa(r = big[,-c(1,2,ncol(big))],nfactors = 5,rotate = "varimax",fm = "ml")   
fa.diagram(fit,main = "diagram illustrating the factors structure")



The theory says that the factors are uncorrelated so , we used *Orthogonal (varimax)* rotation which prohibit the factors from correlations , and as the data is normal , so we used *maximum likelihood* Extraction method

fit$loadings%>%matrix(ncol = 5,dimnames = list(c(names(big[,-c(1,2,ncol(big))])),paste0("F",1:5)))%>%kable(format = "html",digits = 3,caption = "Factor Items loading",row.names = T)

Factor Items loading

F1

F2

F3

F4

F5

E1

0.632

0.073

-0.064

0.054

0.003

E2

0.567

0.071

-0.109

0.101

0.066

E3

0.667

0.193

-0.109

0.031

0.131

E4

0.511

-0.017

0.001

0.128

0.149

E6

0.563

0.047

-0.380

0.011

-0.140

E7

0.631

0.033

-0.114

0.131

0.166

N1

-0.011

-0.102

0.556

0.063

-0.100

N2

-0.226

0.035

0.611

0.103

-0.027

N3

-0.412

0.005

0.580

0.115

0.041

N5

-0.061

-0.030

0.726

-0.046

-0.084

N8

0.078

-0.104

0.609

-0.012

-0.071

N9

-0.344

-0.050

0.511

0.049

0.085

A1

-0.045

0.247

-0.005

0.533

0.078

A2

0.061

0.026

0.020

0.576

0.049

A3

0.042

0.093

0.088

0.672

0.105

A4

0.141

0.087

0.016

0.619

0.132

A5

0.277

0.215

0.110

0.603

0.195

A6

0.181

0.278

0.013

0.543

0.126

O2

0.032

0.187

-0.037

0.024

0.710

O3

0.058

0.117

-0.012

0.169

0.429

O4

0.100

0.196

0.015

0.170

0.631

O5

0.122

0.125

-0.148

0.162

0.644

O6

0.135

0.285

-0.057

0.051

0.416

O8

0.363

0.142

0.024

0.229

0.192

C3

0.120

0.524

0.116

0.131

0.285

C4

0.257

0.405

-0.089

0.106

0.279

C5

0.240

0.531

-0.062

0.161

0.222

C6

-0.045

0.802

-0.089

0.181

0.076

C7

0.077

0.721

-0.079

0.234

0.192

C8

0.025

0.453

-0.209

0.140

0.218

fit$r.scores%>%kable(format = "html",digits = 3,caption = "factor correlation")

factor correlation

ML3

ML1

ML2

ML5

ML4

ML3

1.000

0.011

-0.083

0.053

0.034

ML1

0.011

1.000

-0.033

0.091

0.096

ML2

-0.083

-0.033

1.000

0.054

-0.006

ML5

0.053

0.091

0.054

1.000

0.061

ML4

0.034

0.096

-0.006

0.061

1.000