This is an R HTML document. When you click the **Knit HTML** button a web page will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
# Factor Analysis
#Exploratory Factor Analysis (EFA) is a statistical technique that is used to identify
#the latent relational structure among a set of variables and narrow down to smaller number
#of variables. This essentially means that the variance of large number of variables can be
#described by few summary variables, i.e., factors.
# Importing the dataset

##############reading the dataset###############################

who <- read.csv("C:/Users/prera/Downloads/Life_Expectancy_Data.csv")
head(who)
```

```
##       Country Year     Status Life.expectancy Adult.Mortality infant.deaths
## 1 Afghanistan 2015 Developing            65.0             263            62
## 2 Afghanistan 2014 Developing            59.9             271            64
## 3 Afghanistan 2013 Developing            59.9             268            66
## 4 Afghanistan 2012 Developing            59.5             272            69
## 5 Afghanistan 2011 Developing            59.2             275            71
## 6 Afghanistan 2010 Developing            58.8             279            74
##   Alcohol percentage.expenditure Hepatitis.B Measles  BMI under.five.deaths
## 1    0.01              71.279624          65    1154 19.1                83
## 2    0.01              73.523582          62     492 18.6                86
## 3    0.01              73.219243          64     430 18.1                89
## 4    0.01              78.184215          67    2787 17.6                93
## 5    0.01               7.097109          68    3013 17.2                97
## 6    0.01              79.679367          66    1989 16.7               102
##   Polio Total.expenditure Diphtheria HIV.AIDS       GDP Population
## 1     6              8.16         65      0.1 584.25921   33736494
## 2    58              8.18         62      0.1 612.69651     327582
## 3    62              8.13         64      0.1 631.74498   31731688
## 4    67              8.52         67      0.1 669.95900    3696958
## 5    68              7.87         68      0.1  63.53723    2978599
## 6    66              9.20         66      0.1 553.32894    2883167
##   thinness..1.19.years thinness.5.9.years Income.composition.of.resources
## 1                 17.2               17.3                           0.479
## 2                 17.5               17.5                           0.476
## 3                 17.7               17.7                           0.470
## 4                 17.9               18.0                           0.463
## 5                 18.2               18.2                           0.454
## 6                 18.4               18.4                           0.448
##   Schooling
## 1      10.1
## 2      10.0
## 3       9.9
## 4       9.8
## 5       9.5
## 6       9.2
```

```
#############Dimesion of the dataset############################

dim(who)
```

```
## [1] 2938   22
```

```
##############      TOP 10 DEVELPED & DEVELOPING Countires       #####################

status.of.countries <- who[(who$Status %in% c("Developing") & who$Life.expectancy<55) | (who$Status %in% c("Developed") & who$Life.expectancy>80) ,]
dim(status.of.countries)
```

```
## [1] 509  22
```

```
#View(status.of.countries)

class(status.of.countries)
```

```
## [1] "data.frame"
```

```
head(status.of.countries)
```

```
##        Country Year     Status Life.expectancy Adult.Mortality infant.deaths
## 16 Afghanistan 2000 Developing            54.8             321            88
## 49      Angola 2015 Developing            52.4             335            66
## 50      Angola 2014 Developing            51.7             348            67
## 51      Angola 2013 Developing            51.1             355            69
## 53      Angola 2011 Developing            51.0             361            75
## 54      Angola 2010 Developing            49.6             365            78
##    Alcohol percentage.expenditure Hepatitis.B Measles  BMI under.five.deaths
## 16    0.01               10.42496          62    6532 12.2               122
## 49      NA                0.00000          64     118 23.3                98
## 50    8.33               23.96561          64   11699 22.7               101
## 51    8.10               35.95857          77    8523 22.1               105
## 53    8.06              239.89139          72    1449 21.0               115
## 54    7.80              191.65374          77    1190  2.4               121
##    Polio Total.expenditure Diphtheria HIV.AIDS       GDP Population
## 16    24              8.20         24      0.1  114.5600     293756
```

```
## 49     7                NA         64      1.9 3695.7937    2785935
## 50    68              3.31         64      2.0  479.3122    2692466
## 51    67              4.26         77      2.3  484.6169    2599834
## 53    73              3.38         71      2.5 4299.1289   24218565
## 54    81              3.39         77      2.5 3529.5348   23369131
##    thinness..1.19.years thinness.5.9.years Income.composition.of.resources
## 16                  2.3                2.5                           0.338
## 49                  8.3                8.2                           0.531
## 50                  8.5                8.3                           0.527
## 51                  8.6                8.5                           0.523
## 53                  8.9                8.8                           0.495
## 54                  9.1                9.0                           0.488
##    Schooling
## 16       5.5
## 49      11.4
## 50      11.4
## 51      11.4
## 53       9.4
## 54       9.0
```

```
#View(status.of.countries)
WHONew<-status.of.countries
#resting the index values
row.names(WHONew) <- NULL
#View(WHONew)
dim(WHONew)
```

```
## [1] 509  22
```

```
###################### CLEANING THE DATA ######################################
# For 347 rows running the for loop for chechking any NA values and replacing it with the mean of the
# particular country.
for(i in 1:347)
{
  if(is.na(WHONew$Alcohol[i]))
  {
    WHONew$Alcohol[i] <- with(WHONew, mean(WHONew$Alcohol[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
for(i in 1:347)
{
  if(is.na(WHONew$Hepatitis.B[i]))
  {
    WHONew$Hepatitis.B[i] <- with(WHONew, mean(WHONew$Hepatitis.B[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
for(i in 1:347)
{
  if(is.na(WHONew$Total.expenditure[i]))
  {
    WHONew$Total.expenditure[i] <- with(WHONew, mean(WHONew$Total.expenditure[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
dim(WHONew)
```

```
## [1] 509  22
```

```
#View(WHONew)

# Deleting the Empty rows where there is no data present.
new.life<- na.omit(WHONew)
dim(new.life)
```

```
## [1] 285  22
```

```
View(new.life)


fac.eff <- new.life[ ,c(4,5,7,11,16,17,19,20,21,22)]

View(fac.eff)
attach(fac.eff)
head(fac.eff[1])
```

```
##   Life.expectancy
## 1            54.8
## 2            52.4
## 3            51.7
## 4            51.1
## 5            51.0
## 6            49.6
```

```
# Computing Correlation Matrix
corrm.fac.eff <- cor(fac.eff[-1])
corrm.fac.eff
```

```
##                                   Adult.Mortality    Alcohol        BMI
## Adult.Mortality                         1.0000000 -0.4530275 -0.4153265
## Alcohol                                -0.4530275  1.0000000  0.5762918
```

```
## BMI                                    -0.4153265  0.5762918  1.0000000
## HIV.AIDS                                0.5676648 -0.3364188 -0.2739551
## GDP                                    -0.4023978  0.5137808  0.4557127
## thinness..1.19.years                    0.5345618 -0.3924699 -0.5408240
## thinness.5.9.years                      0.5538395 -0.3747142 -0.5176442
## Income.composition.of.resources        -0.4775719  0.6980880  0.7668355
## Schooling                              -0.4277133  0.7156089  0.7487213
##                                   HIV.AIDS        GDP thinness..1.19.years
## Adult.Mortality                  0.5676648 -0.4023978            0.5345618
## Alcohol                         -0.3364188  0.5137808           -0.3924699
## BMI                             -0.2739551  0.4557127           -0.5408240
## HIV.AIDS                         1.0000000 -0.3260591            0.4341843
## GDP                             -0.3260591  1.0000000           -0.4571052
## thinness..1.19.years             0.4341843 -0.4571052            1.0000000
## thinness.5.9.years               0.4631060 -0.4526723            0.9188703
## Income.composition.of.resources -0.3450488  0.6115709           -0.6638266
## Schooling                       -0.2729460  0.5859284           -0.5865751
##                                 thinness.5.9.years
## Adult.Mortality                          0.5538395
## Alcohol                                 -0.3747142
## BMI                                     -0.5176442
## HIV.AIDS                                 0.4631060
## GDP                                     -0.4526723
## thinness..1.19.years                     0.9188703
## thinness.5.9.years                       1.0000000
## Income.composition.of.resources         -0.6422292
## Schooling                               -0.5678306
##                                 Income.composition.of.resources  Schooling
## Adult.Mortality                                      -0.4775719 -0.4277133
## Alcohol                                               0.6980880  0.7156089
## BMI                                                   0.7668355  0.7487213
## HIV.AIDS                                             -0.3450488 -0.2729460
## GDP                                                   0.6115709  0.5859284
## thinness..1.19.years                                 -0.6638266 -0.5865751
## thinness.5.9.years                                   -0.6422292 -0.5678306
## Income.composition.of.resources                       1.0000000  0.9347573
## Schooling                                             0.9347573  1.0000000
```

You can also embed plots, for example:

```
install.packages("corrplot")
```

```
## Installing package into 'C:/Users/prera/OneDrive/Documents/R/win-library/3.6'
## (as 'lib' is unspecified)
```

```
## Error in contrib.url(repos, "source"): trying to use CRAN without setting a mirror
```

```
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 3.6.3
```

```
## corrplot 0.84 loaded
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#For plotting the scatter density plots
library(GGally)
```

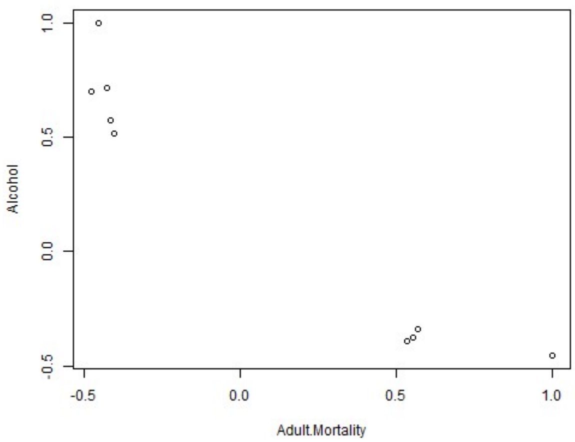
```
## Loading required package: ggplot2
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
##
## Attaching package: 'GGally'
```

```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

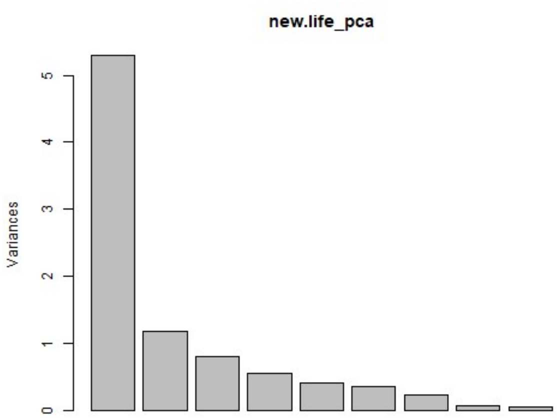
```r
plot(corrm.fac.eff)
```




```r
new.life_pca <- prcomp(fac.eff[-1], scale=TRUE)
summary(new.life_pca)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.3031 1.0868 0.89854 0.75009 0.64178 0.59893 0.48578
## Proportion of Variance 0.5894 0.1312 0.08971 0.06252 0.04576 0.03986 0.02622
## Cumulative Proportion  0.5894 0.7206 0.81032 0.87283 0.91860 0.95845 0.98467
##                            PC8     PC9
## Standard deviation     0.28324 0.24023
## Proportion of Variance 0.00891 0.00641
## Cumulative Proportion  0.99359 1.00000
```

```r
plot(new.life_pca)
```




```r
# A table containing eigenvalues and %'s accounted, follows. Eigenvalues are the sdev^2
(eigen.fac.eff <- round(new.life_pca$sdev^2,2))
```

```
## [1] 5.30 1.18 0.81 0.56 0.41 0.36 0.24 0.08 0.06
```

```r
names(eigen.fac.eff) <- paste("PC",1:9,sep="")
eigen.fac.eff
```

```
##  PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9
## 5.30 1.18 0.81 0.56 0.41 0.36 0.24 0.08 0.06
```

```r
sumlambdas <- sum(eigen.fac.eff)
sumlambdas
```

```
## [1] 9
```

```
cumvar.fac.eff <- cumsum(propvar)
```

```
## Error in eval(expr, envir, enclos): object 'propvar' not found
```

```
propvar <- round(eigen.fac.eff/sumlambdas,2)
propvar
```

```
##  PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9
## 0.59 0.13 0.09 0.06 0.05 0.04 0.03 0.01 0.01
```

```
cumvar.fac.eff <- cumsum(propvar)
cumvar.fac.eff
```

```
##  PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9
## 0.59 0.72 0.81 0.87 0.92 0.96 0.99 1.00 1.01
```

```
matlambdas <- rbind(eigen.fac.eff,propvar,cumvar.fac.eff)
matlambdas
```

```
##                 PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9
## eigen.fac.eff  5.30 1.18 0.81 0.56 0.41 0.36 0.24 0.08 0.06
## propvar        0.59 0.13 0.09 0.06 0.05 0.04 0.03 0.01 0.01
## cumvar.fac.eff 0.59 0.72 0.81 0.87 0.92 0.96 0.99 1.00 1.01
```

```
rownames(matlambdas) <- c("Eigenvalues","Prop. variance","Cum. prop. variance")
rownames(matlambdas)
```

```
## [1] "Eigenvalues"         "Prop. variance"      "Cum. prop. variance"
```

```
eigvec.emp <- new.life_pca$rotation
print(new.life_pca)
```

```
## Standard deviations (1, .., p=9):
## [1] 2.3031386 1.0867502 0.8985405 0.7500902 0.6417774 0.5989313 0.4857819
## [8] 0.2832368 0.2402331
##
## Rotation (n x k) = (9 x 9):
##                                      PC1       PC2         PC3         PC4
## Adult.Mortality                 -0.2963146 0.3933735  0.32256464 -0.18714120
## Alcohol                          0.3209942 0.2975146 -0.41755945  0.16611060
## BMI                              0.3411230 0.2733771  0.09577648  0.36221151
## HIV.AIDS                        -0.2374740 0.5433865  0.46615149 -0.08782235
## GDP                              0.3019017 0.1277741 -0.21359316 -0.88233421
## thinness..1.19.years            -0.3519585 0.2828611 -0.48138414  0.05968413
## thinness.5.9.years              -0.3485108 0.3260066 -0.45961828  0.06859044
## Income.composition.of.resources  0.3960508 0.2543053  0.07153997  0.06224593
## Schooling                        0.3782407 0.3456159  0.03749890  0.08843326
##                                        PC5         PC6         PC7
## Adult.Mortality                  0.739488143 -0.21630783  0.15597097
## Alcohol                         -0.084880468 -0.62808722  0.45128571
## BMI                              0.133005364  0.65476669  0.46983943
## HIV.AIDS                        -0.642072242 -0.04957061  0.08119827
## GDP                             -0.039747971  0.22195785  0.13095165
## thinness..1.19.years             0.005153132  0.19628897 -0.16025755
## thinness.5.9.years               0.006619857  0.18872320 -0.12723461
## Income.composition.of.resources  0.099567043 -0.01561493 -0.45582342
## Schooling                        0.066644790 -0.06233604 -0.52714093
##                                        PC8          PC9
## Adult.Mortality                 -0.0204520520 -0.002848616
## Alcohol                         -0.0141814025 -0.030675532
## BMI                             -0.0246304217  0.027643013
## HIV.AIDS                        -0.0186365347 -0.040898855
## GDP                             -0.0003984935  0.019994326
## thinness..1.19.years            -0.6916551644 -0.134895083
## thinness.5.9.years               0.7075942789  0.061381753
## Income.composition.of.resources  0.1034032951 -0.735428718
## Schooling                       -0.0930467252  0.658327024
```

```
#Considering the observation of the standard deviation, we initially tried considering four
#components as the SD after four components are comparatively higher also considering the
#elbow diagram done in the previous section justifies for taking into account four components.
# Taking the first four PCs to generate linear combinations for all the variables with four factors
pcafactors.emp <- eigvec.emp[,1:4]
pcafactors.emp
```

```
##                                      PC1       PC2         PC3         PC4
## Adult.Mortality                 -0.2963146 0.3933735  0.32256464 -0.18714120
## Alcohol                          0.3209942 0.2975146 -0.41755945  0.16611060
## BMI                              0.3411230 0.2733771  0.09577648  0.36221151
## HIV.AIDS                        -0.2374740 0.5433865  0.46615149 -0.08782235
## GDP                              0.3019017 0.1277741 -0.21359316 -0.88233421
## thinness..1.19.years            -0.3519585 0.2828611 -0.48138414  0.05968413
## thinness.5.9.years              -0.3485108 0.3260066 -0.45961828  0.06859044
## Income.composition.of.resources  0.3960508 0.2543053  0.07153997  0.06224593
## Schooling                        0.3782407 0.3456159  0.03749890  0.08843326
```

```r
# Multiplying each column of the eigenvector's matrix by the square-root of the corresponding eigenvalue in order to get the factor loadings
unrot.fact.emp <- sweep(pcafactors.emp,MARGIN=2,new.life_pca$sdev[1:4],`*`)
unrot.fact.emp
```

```
##                                      PC1       PC2         PC3         PC4
## Adult.Mortality                 -0.6824537 0.4274988  0.28983739 -0.14037279
## Alcohol                          0.7392942 0.3233240 -0.37519408  0.12459794
## BMI                              0.7856536 0.2970926  0.08605905  0.27169132
## HIV.AIDS                        -0.5469356 0.5905254  0.41885599 -0.06587469
## GDP                              0.6953215 0.1388585 -0.19192210 -0.66183028
## thinness..1.19.years            -0.8106092 0.3073993 -0.43254314  0.04476849
## thinness.5.9.years              -0.8026688 0.3542877 -0.41298564  0.05144902
## Income.composition.of.resources  0.9121599 0.2763663  0.06428156  0.04669007
## Schooling                        0.8711407 0.3755981  0.03369428  0.06633293
```

```r
# Computing communalities
communalities.emp <- rowSums(unrot.fact.emp^2)
communalities.emp
```

```
##                 Adult.Mortality                         Alcohol
##                       0.7522084                       0.8073897
##                             BMI                        HIV.AIDS
##                       0.7867380                       0.8276387
##                             GDP            thinness..1.19.years
##                       0.9776071                       0.9406794
##              thinness.5.9.years Income.composition.of.resources
##                       0.9430011                       0.9147262
##                       Schooling
##                       0.9054955
```

```r
# Performing the varimax rotation. The default in the varimax function is norm=TRUE thus, Kaiser normalization is carried out
rot.fact.emp <- varimax(unrot.fact.emp)
View(unrot.fact.emp)
rot.fact.emp
```

```
## $loadings
##
## Loadings:
##                                 PC1    PC2    PC3    PC4
## Adult.Mortality                 -0.288  0.759 -0.293
## Alcohol                          0.795 -0.330        -0.255
## BMI                              0.812 -0.116  0.337
## HIV.AIDS                                0.877 -0.194  0.123
## GDP                              0.340 -0.179  0.213 -0.886
## thinness..1.19.years            -0.299  0.261 -0.872  0.150
## thinness.5.9.years              -0.265  0.300 -0.872  0.151
## Income.composition.of.resources  0.805 -0.152  0.411 -0.271
## Schooling                        0.849         0.327 -0.262
##
##                  PC1   PC2   PC3   PC4
## SS loadings    3.026 1.688 2.080 1.062
## Proportion Var 0.336 0.188 0.231 0.118
## Cumulative Var 0.336 0.524 0.755 0.873
##
## $rotmat
##            [,1]       [,2]        [,3]       [,4]
## [1,]  0.6848287 -0.4010275  0.52019938 -0.3155618
## [2,]  0.6184353  0.6653274 -0.39108202 -0.1480948
## [3,] -0.1140636  0.6018484  0.75535942  0.2328092
## [4,]  0.3681535 -0.1851886 -0.07667864  0.9079034
```

```r
# The print method of varimax omits loadings less than abs(0.1). In order to display all the loadings, it is necessary to ask explicitly the contents
fact.load.emp <- rot.fact.emp$loadings[1:9,1:4]
fact.load.emp
```

```
##                                         PC1         PC2         PC3         PC4
## Adult.Mortality                 -0.28772215  0.75854298 -0.29250406  0.09207786
## Alcohol                          0.79491204 -0.33024507 -0.03482619 -0.25540135
## BMI                              0.81197857 -0.11592453  0.33668157 -0.02521530
## HIV.AIDS                        -0.08138367  0.87651605 -0.19402144  0.12284395
## GDP                              0.34028725 -0.17940131  0.21317882 -0.88554030
## thinness..1.19.years            -0.29920268  0.26098181 -0.87205507  0.15021846
## thinness.5.9.years              -0.26453881  0.29952710 -0.87200099  0.15138731
## Income.composition.of.resources  0.80544487 -0.15188587  0.41139868 -0.27141580
## Schooling                        0.84944267 -0.09146092  0.32664215 -0.26245460
```

```r
scale.emp <- scale(fac.eff[-1])
head(scale.emp)
```

```
##   Adult.Mortality     Alcohol        BMI   HIV.AIDS        GDP
## 1       0.3025231 -1.46317163 -0.9386706 -0.7355316 -0.4957402
## 2       0.3661937 -0.04793103 -0.3991077 -0.5888599 -0.3173613
## 3       0.4253164  0.65759545 -0.4282733 -0.5807114 -0.4775722
## 4       0.4571517  0.59896848 -0.4574388 -0.5562661 -0.4773079
## 5       0.4844391  0.58877248 -0.5109090 -0.5399693 -0.2873095
## 6       0.5026307  0.52249851 -1.4150414 -0.5399693 -0.3256425
##   thinness..1.19.years thinness.5.9.years Income.composition.of.resources
## 1           -0.6650011         -0.6259672                      -0.8934474
```

```
## 2           0.8032389          0.7360523                          -0.1276533
## 3           0.8521802          0.7599473                          -0.1435246
## 4           0.8766509          0.8077375                          -0.1593960
## 5           0.9500629          0.8794227                          -0.2704957
## 6           0.9990042          0.9272129                          -0.2982706
##     Schooling
## 1 -1.30128198
## 2  0.03496574
## 3  0.03496574
## 4  0.03496574
## 5 -0.41799959
## 6 -0.50859265
```

```
#as.matrix(scale.emp)%%fact.load.emp%%solve(t(fact.load.emp)%*%fact.load.emp)

library(psych)
```

```
## Warning: package 'psych' was built under R version 3.6.3
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
#install.packages("psych", lib="/Library/Frameworks/R.framework/Versions/3.5/Resources/library")
library(psych)
# On 4 factors analysis
fit.pc <- principal(fac.eff[-1], nfactors=4, rotate="varimax")
# On 3 factors analysis
fit.pc1 <- principal(fac.eff[-1], nfactors=3, rotate="varimax")

fit.pc
```

```
## Principal Components Analysis
## Call: principal(r = fac.eff[-1], nfactors = 4, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                                   RC1   RC3   RC2   RC4   h2    u2 com
## Adult.Mortality                 -0.29  0.29  0.76 -0.09 0.75 0.248 1.6
## Alcohol                          0.79  0.03 -0.33  0.26 0.81 0.193 1.6
## BMI                              0.81 -0.34 -0.12  0.03 0.79 0.213 1.4
## HIV.AIDS                        -0.08  0.19  0.88 -0.12 0.83 0.172 1.2
## GDP                              0.34 -0.21 -0.18  0.89 0.98 0.022 1.5
## thinness..1.19.years            -0.30  0.87  0.26 -0.15 0.94 0.059 1.5
## thinness.5.9.years              -0.26  0.87  0.30 -0.15 0.94 0.057 1.5
## Income.composition.of.resources  0.81 -0.41 -0.15  0.27 0.91 0.085 1.8
## Schooling                        0.85 -0.33 -0.09  0.26 0.91 0.095 1.5
##
##                        RC1  RC3  RC2  RC4
## SS loadings           3.03 2.08 1.69 1.06
## Proportion Var        0.34 0.23 0.19 0.12
## Cumulative Var        0.34 0.57 0.75 0.87
## Proportion Explained  0.39 0.26 0.21 0.14
## Cumulative Proportion 0.39 0.65 0.86 1.00
##
## Mean item complexity =  1.5
## Test of the hypothesis that 4 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.05
##  with the empirical chi square  47.49  with prob <  1.5e-08
##
## Fit based upon off diagonal values = 0.99
```

```
fit.pc1
```

```
## Principal Components Analysis
## Call: principal(r = fac.eff[-1], nfactors = 3, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                                   RC1   RC3   RC2   h2    u2 com
## Adult.Mortality                 -0.29  0.30  0.75 0.73 0.267 1.7
## Alcohol                          0.83  0.01 -0.31 0.79 0.208 1.3
## BMI                              0.75 -0.38 -0.07 0.71 0.287 1.5
## HIV.AIDS                        -0.12  0.19  0.88 0.82 0.177 1.1
## GDP                              0.64 -0.18 -0.31 0.54 0.460 1.6
## thinness..1.19.years            -0.31  0.88  0.27 0.94 0.061 1.5
## thinness.5.9.years              -0.28  0.87  0.31 0.94 0.060 1.5
## Income.composition.of.resources  0.84 -0.44 -0.15 0.91 0.087 1.6
## Schooling                        0.88 -0.35 -0.08 0.90 0.099 1.3
##
##                        RC1  RC3  RC2
## SS loadings           3.42 2.15 1.73
## Proportion Var        0.38 0.24 0.19
## Cumulative Var        0.38 0.62 0.81
## Proportion Explained  0.47 0.29 0.24
## Cumulative Proportion 0.47 0.76 1.00
##
## Mean item complexity =  1.4
## Test of the hypothesis that 3 components are sufficient.
##
```

```
## The root mean square of the residuals (RMSR) is  0.05
##  with the empirical chi square  61.3  with prob <  1.3e-08
##
## Fit based upon off diagonal values = 0.99
```

```
round(fit.pc$values, 3)
```

```
## [1] 5.304 1.181 0.807 0.563 0.412 0.359 0.236 0.080 0.058
```

```
round(fit.pc1$values, 3)
```

```
## [1] 5.304 1.181 0.807 0.563 0.412 0.359 0.236 0.080 0.058
```

```
fit.pc$loadings
```

```
##
## Loadings:
##                                RC1    RC3    RC2    RC4
## Adult.Mortality                -0.288  0.293  0.759
## Alcohol                         0.795        -0.330  0.255
## BMI                             0.812 -0.337 -0.116
## HIV.AIDS                               0.194  0.877 -0.123
## GDP                             0.340 -0.213 -0.179  0.886
## thinness..1.19.years           -0.299  0.872  0.261 -0.150
## thinness.5.9.years             -0.265  0.872  0.300 -0.151
## Income.composition.of.resources  0.805 -0.411 -0.152  0.271
## Schooling                       0.849 -0.327         0.262
##
##                  RC1   RC3   RC2   RC4
## SS loadings    3.026 2.080 1.688 1.062
## Proportion Var 0.336 0.231 0.188 0.118
## Cumulative Var 0.336 0.567 0.755 0.873
```

```
fit.pc1$loadings
```

```
##
## Loadings:
##                                RC1    RC3    RC2
## Adult.Mortality                -0.293  0.303  0.745
## Alcohol                         0.833        -0.314
## BMI                             0.753 -0.377
## HIV.AIDS                       -0.115  0.192  0.879
## GDP                             0.643 -0.176 -0.308
## thinness..1.19.years           -0.311  0.876  0.272
## thinness.5.9.years             -0.280  0.874  0.312
## Income.composition.of.resources  0.838 -0.436 -0.146
## Schooling                       0.877 -0.354
##
##                  RC1   RC3   RC2
## SS loadings    3.419 2.149 1.725
## Proportion Var 0.380 0.239 0.192
## Cumulative Var 0.380 0.619 0.810
```

```
# Loadings with more digits
for (i in c(1,3,2,4)) { print(fit.pc$loadings[[1,i]])}  # 4 components
```

```
## [1] -0.2877222
## [1] 0.758543
## [1] 0.2925041
## [1] -0.09207786
```

```
for (i in c(1,3,2,4)) { print(fit.pc1$loadings[[1,i]])} # 3 components
```

```
## [1] -0.2925753
## [1] 0.745044
## [1] 0.3030066
```

```
## Error in fit.pc1$loadings[[1, i]]: subscript out of bounds
```

```
# Communalities
fit.pc$communality
```

```
##                 Adult.Mortality                         Alcohol
##                       0.7522084                       0.8073897
##                             BMI                        HIV.AIDS
##                       0.7867380                       0.8276387
##                             GDP            thinness..1.19.years
##                       0.9776071                       0.9406794
##              thinness.5.9.years Income.composition.of.resources
##                       0.9430011                       0.9147262
##                       Schooling
##                       0.9054955
```

```
fit.pc1$communality
```

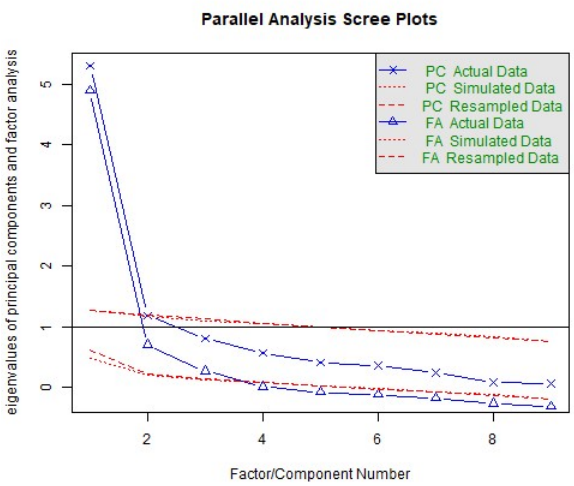
```
##                Adult.Mortality                        Alcohol
##                      0.7325039                      0.7918650
##                            BMI                       HIV.AIDS
##                      0.7129218                      0.8232992
##                            GDP            thinness..1.19.years
##                      0.5395878                      0.9386752
##             thinness.5.9.years Income.composition.of.resources
##                      0.9403541                      0.9125462
##                      Schooling
##                      0.9010954
```

```
# Rotated factor scores, Notice the columns ordering: RC1, RC3, RC2 and RC4
head(fit.pc$scores)
```

```
##           RC1        RC3        RC2         RC4
## 1 -1.87231179 -1.2526059 -0.3875991 -0.21184580
## 2  0.03201845  0.9441829 -0.4093176 -0.12371144
## 3  0.34852090  1.2559007 -0.5349032 -0.29552668
## 4  0.32354251  1.2701292 -0.4984147 -0.27534799
## 5  0.08707079  1.3592718 -0.5480157 -0.01157608
## 6 -0.33601387  1.4056528 -0.6285787  0.29479789
```

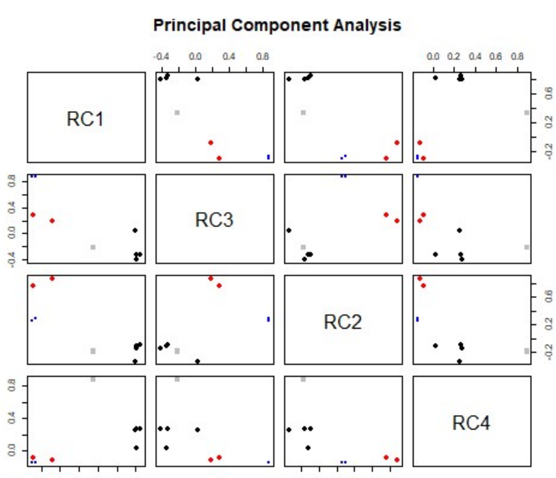
```
# Play with FA utilities

fa.parallel(fac.eff[-1]) # See factor recommendation
```

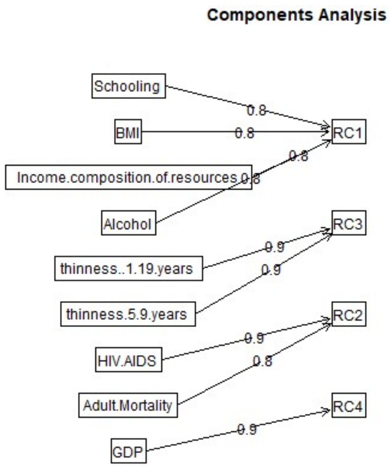



```
## Parallel analysis suggests that the number of factors =  3  and the number of components =  1
```

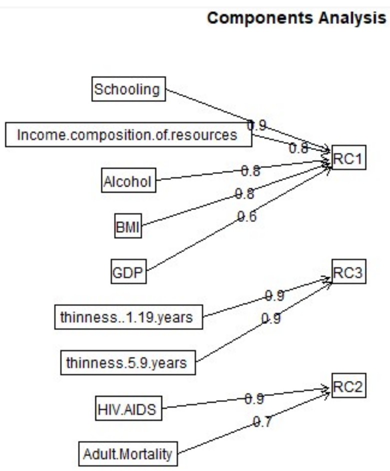
```
fa.plot(fit.pc) # See Correlations within Factors
```




```
fa.diagram(fit.pc)
```



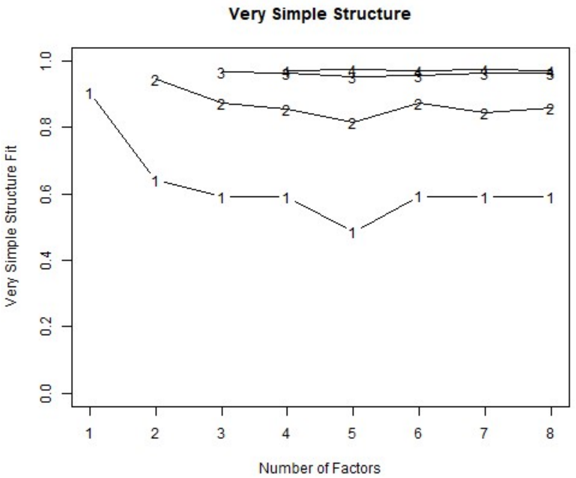

```r
fa.diagram(fit.pc1)# Visualize the relationship for 4 components analysis
```




```r
#We have observed 9 variables schooling, BMI, income composition of resources, Alcohol,
#thinness 5.9 years, thinness 1.91 years, HIV/AIDS, Adult Mortality, GDP. We hypothesize
#that there are four unobserved latent factors (RC1, RC2, RC3, RC4) that underly the observed
#variables as described in this diagram. Schooling, BMI, income composition of resources,
#Alcohol loads on RC1 with loadings 0.8, 0.8, 0.8 and 0.9 respectively. Thinness 5.9 years,
#thinness 1.91 years loads on RC2 with loadings 0.9 and 0.9. HIV/AIDS, Adult Mortality on
#RC3 with loadings 0.9 and 0.8.GDP loads on RC4 with loadings of 0.9.

#In a three-component analysis, we see that the common variance for the GDP and adult mortality
#is comparatively less considering the four component analysis for the same. As the unique
#variance of 0.4 GDP and unique variance of 0.3 Adult mortality is given away in 3 component
#analysis. Where as, in 4 component analysis we obsevre the unique vareince of GDP, Adult Mortality
#is only 0.1 and 0.2 respectievly is given away. Therefore, we take Component as 4 for analysis.

vss(fac.eff[-1]) # See Factor recommendations for a simple structure
```




```
##
## Very Simple Structure
## Call: vss(x = fac.eff[-1])
## VSS complexity 1 achieves a maximimum of 0.9  with  1  factors
## VSS complexity 2 achieves a maximimum of 0.94  with  2  factors
##
## The Velicer MAP achieves a minimum of 0.08  with  3  factors
## BIC achieves a minimum of  NA  with  3  factors
## Sample Size adjusted BIC achieves a minimum of  NA  with  3  factors
##
## Statistics by number of factors
##   vss1 vss2   map dof   chisq     prob sqresid  fit RMSEA   BIC SABIC complex
## 1 0.90 0.00 0.110  27 6.7e+02 1.4e-123    2.95 0.90 0.289 516.0 601.6     1.0
## 2 0.64 0.94 0.091  19 1.6e+02  1.6e-24    1.73 0.94 0.162  53.7 113.9     1.4
## 3 0.59 0.87 0.084  12 1.0e+01  6.1e-01    1.02 0.97 0.000 -57.8 -19.7     1.5
## 4 0.59 0.86 0.140   6 7.3e+00  2.9e-01    0.89 0.97 0.027 -26.6  -7.6     1.6
## 5 0.49 0.82 0.257   1 8.6e-01  3.5e-01    0.74 0.98 0.000  -4.8  -1.6     1.7
## 6 0.59 0.87 0.508  -3 1.1e-05       NA    0.86 0.97    NA    NA    NA     1.6
## 7 0.59 0.84 0.448  -6 1.5e-07       NA    0.83 0.97    NA    NA    NA     1.7
## 8 0.59 0.86 1.000  -8 0.0e+00       NA    0.83 0.97    NA    NA    NA     1.7
##    eChisq    SRMR  eCRMS  eBIC
## 1 2.3e+02 1.1e-01 0.1215  74.5
## 2 5.7e+01 5.3e-02 0.0724 -50.6
## 3 1.5e+00 8.5e-03 0.0148 -66.3
## 4 3.1e-01 3.9e-03 0.0095 -33.6
## 5 9.4e-03 6.8e-04 0.0041  -5.6
## 6 8.1e-07 6.3e-06     NA    NA
## 7 1.6e-08 8.9e-07     NA    NA
## 8 2.5e-17 3.5e-11     NA    NA
```