```
R <- cor(lifeex[,-1])
> R
           0m       25m       50m       75m        0w       25w       50w
75w
0m  1.0000000 0.7483563 0.6356374 0.2898761 0.9801778 0.8739607 0.6965615 0.3
175692
25m 0.7483563 1.0000000 0.6670171 0.3911019 0.6933171 0.7247417 0.6474890 0.3
930184
50m 0.6356374 0.6670171 1.0000000 0.7519835 0.5574436 0.7716579 0.8021425 0.5
931704
75m 0.2898761 0.3911019 0.7519835 1.0000000 0.2472239 0.5466748 0.6869444 0.7
104841
0w  0.9801778 0.6933171 0.5574436 0.2472239 1.0000000 0.8873568 0.7098933 0.3
651831
25w 0.8739607 0.7247417 0.7716579 0.5466748 0.8873568 1.0000000 0.9399013 0.6
843048
50w 0.6965615 0.6474890 0.8021425 0.6869444 0.7098933 0.9399013 1.0000000 0.8
279568
75w 0.3175692 0.3930184 0.5931704 0.7104841 0.3651831 0.6843048 0.8279568 1.0
000000
```

Life expectancy for newborn males and females have a very high correlation of 98%. The life expectancy of 25 year old women and 50 year old women also have a very high correlation of 94%. 50 year old women and 75 year old women have a very high correlation in life expectancy of ~83%. 25 year old women and new born men have a very high correlation in life expectancy of ~87%. 25 year old women and new born women have a very high correlation in life expectancy of ~89%.

```
> model.mle1 <- factanal(lifeex[,-1], factors = 1, rotation = "varimax")
> model.mle1

Call:
factanal(x = lifeex[, -1], factors = 1, rotation = "varimax")

Uniquenesses:
   0m   25m   50m   75m    0w   25w   50w   75w
0.238 0.470 0.399 0.696 0.217 0.005 0.117 0.532

Loadings:
    Factor1
0m  0.873
25m 0.728
50m 0.776
75m 0.552
0w  0.885
25w 0.998
50w 0.940
75w 0.684

               Factor1
SS loadings      5.329
Proportion Var   0.666

Test of the hypothesis that 1 factor is sufficient.
The chi square statistic is 163.11 on 20 degrees of freedom.
The p-value is 1.88e-24
```

The test of the hypothesis that 1 factor is sufficient is rejected since the p value is a lot lower than 0.05. Thus 1 factor is not a good model for the data.

```
> model.mle2 <- factanal(lifeex[,-1], factors = 2, rotation = "varimax")
> model.mle2

Call:
factanal(x = lifeex[, -1], factors = 2, rotation = "varimax")

Uniquenesses:
   0m   25m   50m   75m    0w   25w   50w   75w
0.024 0.442 0.346 0.408 0.015 0.011 0.015 0.178

Loadings:
    Factor1 Factor2
0m  0.972   0.179
25m 0.670   0.329
50m 0.480   0.651
75m 0.122   0.760
0w  0.973   0.194
25w 0.790   0.603
50w 0.567   0.815
75w 0.185   0.888

                Factor1 Factor2
SS loadings       3.567   2.994
Proportion Var    0.446   0.374
Cumulative Var    0.446   0.820

Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 45.24 on 13 degrees of freedom.
The p-value is 1.91e-05
> model.mle3 <- factanal(lifeex[,-1], factors = 3, rotation = "varimax")
> model.mle3

Call:
factanal(x = lifeex[, -1], factors = 3, rotation = "varimax")

Uniquenesses:
   0m   25m   50m   75m    0w   25w   50w   75w
0.005 0.362 0.066 0.288 0.005 0.011 0.020 0.146

Loadings:
    Factor1 Factor2 Factor3
0m  0.964   0.122   0.226
25m 0.646   0.169   0.438
50m 0.430   0.354   0.790
75m         0.525   0.656
0w  0.970   0.217
25w 0.764   0.556   0.310
50w 0.536   0.729   0.401
75w 0.156   0.867   0.280

                Factor1 Factor2 Factor3
SS loadings       3.375   2.082   1.640
Proportion Var    0.422   0.260   0.205
Cumulative Var    0.422   0.682   0.887

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 6.73 on 7 degrees of freedom.
The p-value is 0.458
```

The p value for the 3 factor model is 0.458 while the p value for the 2 factor model is very low. Therefore the 3 factor model is the model which fits the data well.

It appears the first factor measures the life expectancy of newborns. The second factor measures life expectancy of older women. The third factor seems to measure the life expectancy of older men.

```
nr.mle3 <- factanal(lifeex[,-1], factors = 3, rotation = "none")
> nr.mle3

Call:
factanal(x = lifeex[, -1], factors = 3, rotation = "none")

Uniquenesses:
   0m   25m   50m   75m    0w   25w   50w   75w
0.005 0.362 0.066 0.288 0.005 0.011 0.020 0.146

Loadings:
    Factor1 Factor2 Factor3
0m   0.982  -0.152
25m  0.748   0.101   0.262
50m  0.680   0.492   0.479
75m  0.376   0.697   0.292
0w   0.984  -0.144
25w  0.943   0.311
50w  0.803   0.577
75w  0.464   0.772  -0.209

               Factor1 Factor2 Factor3
SS loadings      4.843   1.807   0.447
Proportion Var   0.605   0.226   0.056
Cumulative Var   0.605   0.831   0.887

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 6.73 on 7 degrees of freedom.
The p-value is 0.458
> # calculate communalities, specific variances and residuls for no rotat
ion maximum likelihood 3 factors
> L3 <- nr.mle3$load
> # Communalities
> h2 <- apply(L3^2, 1, sum)
> h2
       0m       25m       50m       75m        0w       25w       50w
75w
0.9951744 0.6383053 0.9337235 0.7122469 0.9950975 0.9889328 0.9798796 0.8
540072
> # Specific Variance
> Psi <- nr.mle3$unique
> Psi
        0m        25m        50m        75m         0w        25w
50w
0.00500000 0.36167392 0.06627724 0.28779358 0.00500000 0.01106701 0.02012
006
       75w
0.14597958
> # Residuals
> R-L3%*%t(L3)-diag(Psi)
                0m           25m           50m           75m           0w
0m  -0.0001743520  5.683958e-03 -1.617899e-04  5.071984e-04  1.745355e-04
25m  0.0056839577  2.076217e-05 -1.659723e-02 -3.685255e-02 -5.880846e-03
50m -0.0001617899 -1.659723e-02 -7.702125e-07  1.364936e-02 -3.117193e-05
75m  0.0005071984 -3.685255e-02  1.364936e-02 -4.044213e-05  2.633797e-03
0w   0.0001745355 -5.880846e-03 -3.117193e-05  2.633797e-03 -9.752647e-05
```

```
25w -0.0003404953  1.173959e-03  2.037799e-03 -9.400553e-03  1.892923e-04
50w  0.0005976945  2.168024e-03 -2.836632e-03 -1.723924e-03 -7.840339e-04
75w -0.0017880507  2.278043e-02 -1.915567e-03  5.936939e-02  2.865717e-03
              25w           50w           75w
0m  -3.404953e-04  5.976945e-04 -1.788051e-03
25m  1.173959e-03  2.168024e-03  2.278043e-02
50m  2.037799e-03 -2.836632e-03 -1.915567e-03
75m -9.400553e-03 -1.723924e-03  5.936939e-02
0w   1.892923e-04 -7.840339e-04  2.865717e-03
25w  1.678791e-07  7.546095e-04 -3.635750e-03
50w  7.546095e-04  3.272683e-07 -1.935904e-04
75w -3.635750e-03 -1.935904e-04  1.317901e-05
> # calculate communal
ities, specific varian
ces and residuls for v
arimax maximum likelih
ood 3 factors
> LV3 <- model.mle3$load
> # Communalities
> h2 <- apply(LV3^2, 1, sum)
> h2
       0m        25m        50m        75m         0w        25w        50w
75w
0.9951744 0.6383053 0.9337235 0.7122469 0.9950975 0.9889328 0.9798796 0.8
540072
> # Specific Variance
> Psi <- model.mle3$unique
> Psi
         0m         25m         50m         75m          0w         25w
50w
0.00500000 0.36167392 0.06627724 0.28779358 0.00500000 0.01106701 0.02012
006
       75w
0.14597958
> # Residuals
> R-LV3%*%t(LV3)-diag(Psi)
               0m           25m           50m           75m           0w
0m  -0.0001743520  5.683958e-03 -1.617899e-04  5.071984e-04  1.745355e-04
25m  0.0056839577  2.076217e-05 -1.659723e-02 -3.685255e-02 -5.880846e-03
50m -0.0001617899 -1.659723e-02 -7.702125e-07  1.364936e-02 -3.117193e-05
75m  0.0005071984 -3.685255e-02  1.364936e-02 -4.044213e-05  2.633797e-03
0w   0.0001745355 -5.880846e-03 -3.117193e-05  2.633797e-03 -9.752647e-05
25w -0.0003404953  1.173959e-03  2.037799e-03 -9.400553e-03  1.892923e-04
50w  0.0005976945  2.168024e-03 -2.836632e-03 -1.723924e-03 -7.840339e-04
75w -0.0017880507  2.278043e-02 -1.915567e-03  5.936939e-02  2.865717e-03
              25w           50w           75w
0m  -3.404953e-04  5.976945e-04 -1.788051e-03
25m  1.173959e-03  2.168024e-03  2.278043e-02
50m  2.037799e-03 -2.836632e-03 -1.915567e-03
75m -9.400553e-03 -1.723924e-03  5.936939e-02
0w   1.892923e-04 -7.840339e-04  2.865717e-03
25w  1.678791e-07  7.546095e-04 -3.635750e-03
50w  7.546095e-04  3.272683e-07 -1.935904e-04
75w -3.635750e-03 -1.935904e-04  1.317901e-05
```

The communalities, specific variances and residual matrixes all remained the same regardless of rotation type. However the factors seem to mean something different in the non rotated model. It appears

factor 1 represents young males and young females, factor 2 represents older men and older women and factor 3 looks at elder men and elder women.

```
fit.pc1 <- principal(lifeex[,-1], nfactors=3, rotate="varimax", method="regre
ssion")
> fit.pc1
Principal Components Analysis
Call: principal(r = lifeex[, -1], nfactors = 3, rotate = "varimax",
    method = "regression")
Standardized loadings (pattern matrix) based upon correlation matrix
     RC1  RC2  RC3   h2    u2 com
0m  0.95 0.14 0.21 0.97 0.031 1.1
25m 0.71 0.06 0.56 0.82 0.177 1.9
50m 0.44 0.41 0.74 0.90 0.096 2.2
75m 0.02 0.64 0.71 0.91 0.094 2.0
0w  0.96 0.22 0.07 0.97 0.027 1.1
25w 0.78 0.55 0.26 0.98 0.015 2.0
50w 0.57 0.72 0.34 0.96 0.038 2.4
75w 0.17 0.93 0.22 0.94 0.062 1.2

                       RC1  RC2  RC3
SS loadings           3.50 2.32 1.64
Proportion Var        0.44 0.29 0.20
Cumulative Var        0.44 0.73 0.93
Proportion Explained  0.47 0.31 0.22
Cumulative Proportion 0.47 0.78 1.00

Mean item complexity =  1.7
Test of the hypothesis that 3 components are sufficient.

The root mean square of the residuals (RMSR) is  0.03
 with the empirical chi square  2.05  with prob <  0.96

Fit based upon off diagonal values = 1
```

The standardized factor loadings for the principal component method are very similar to the loadings for the maximum likelihood method of the factor analysis model. Both are very similar.

```
# perform factor analysis for 3 factors using the regression method
> model.reg3 <- factanal(lifeex[,-1], factors = 3, rotation = "varimax", scor
e = "regression")
> model.reg3

Call:
factanal(x = lifeex[, -1], factors = 3, scores = "regression",     rotation =
"varimax")

Uniquenesses:
   0m   25m   50m   75m    0w   25w   50w   75w
0.005 0.362 0.066 0.288 0.005 0.011 0.020 0.146

Loadings:
    Factor1 Factor2 Factor3
0m  0.964   0.122   0.226
25m 0.646   0.169   0.438
50m 0.430   0.354   0.790
75m         0.525   0.656
```

```
0w  0.970   0.217
25w 0.764   0.556   0.310
50w 0.536   0.729   0.401
75w 0.156   0.867   0.280

                Factor1 Factor2 Factor3
SS loadings       3.375   2.082   1.640
Proportion Var    0.422   0.260   0.205
Cumulative Var    0.422   0.682   0.887

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 6.73 on 7 degrees of freedom.
The p-value is 0.458
> model.reg3$scores
            Factor1     Factor2     Factor3
 [1,] -0.258062561  1.90095771  1.91581631
 [2,] -2.782495791 -0.72340014 -1.84772224
 [3,] -2.806428187 -0.81158820 -0.01210318
 [4,]  0.141004934 -0.29028454 -0.85862443
 [5,] -0.196352142  0.47429917 -1.55046466
 [6,]  0.367371307  0.82902375 -0.55214085
 [7,] -1.028567629 -0.08065792 -0.65421971
 [8,]  0.946193522  0.06400408 -0.91995289
 [9,] -0.862493550  3.59177195 -0.36442148
[10,]  1.245304248  0.29564122 -0.27342781
[11,]  0.508736247 -0.50500435  1.01328707
[12,]  0.106044085  0.01111171  1.83871599
[13,] -0.608155779  0.65100820  0.48836431
[14,]  0.235114220 -0.69123901 -0.38558654
[15,]  0.132008172  0.25241049 -0.15220645
[16,] -1.450336359 -0.67765804  0.65911906
[17,]  0.043253249 -1.85175707  0.30633182
[18,]  0.462124701 -0.51918493  0.08032855
[19,] -0.052332675 -0.72020002  0.44417800
[20,]  0.268974443  0.08407227  1.70568388
[21,]  0.442333434 -0.73778272  1.25218728
[22,]  0.711367053 -0.95989475 -0.21545329
[23,]  0.787286051 -1.10729029 -0.51958264
[24,]  1.128331259  0.16389896 -0.68177046
[25,]  0.400058903 -0.36230253 -0.74299137
[26,]  1.214345385  0.40877239 -0.69225320
[27,]  1.128331259  0.16389896 -0.68177046
[28,]  0.731344988  0.24811968 -0.12817725
[29,]  0.009751528  0.75222637 -0.49198911
[30,] -0.240602517 -0.29543613  0.42919600
[31,] -0.723451797  0.44246371  1.59164974



# sort countries by values of first factor scores
> regscfac1 <- model.reg3$scores[,1]
> test1=data.frame(lifeex[ ,1], regscfac1)
> test1[order(test1[ ,2]), ]
          lifeex...1.    regscfac1
3          Madagascar -2.806428187
2            Cameroon -2.782495791
16           Guatemala -1.450336359
7       SouthAfrica(C) -1.028567629
9              Tunisia -0.862493550
31             Ecuador -0.723451797
13          ElSalvador -0.608155779
1              Algeria -0.258062561
30            Columbia -0.240602517
5              Reunion -0.196352142
```

```
19             Mexico -0.052332675
29              Chile  0.009751528
17           Honduras  0.043253249
12       DominicanRep  0.106044085
15            Grenada  0.132008172
4           Mauritius  0.141004934
14          Greenland  0.235114220
20          Nicaragua  0.268974443
6           Seychelles  0.367371307
25 UnitedStates(NW66)  0.400058903
21             Panama  0.442333434
18             Jamaica  0.462124701
11          CostaRica  0.508736247
22       Trinidad(62)  0.711367053
28          Argentina  0.731344988
23       Trinidad(67)  0.787286051
8      SouthAfrica(W)  0.946193522
24   UnitedStates(66)  1.128331259
27   UnitedStates(67)  1.128331259
26  UnitedStates(W66)  1.214345385
10             Canada  1.245304248
> # sort countries by values of second factor scores
> regscfac2 <- model.reg3$scores[,2]
> test2=data.frame(lifeex[ ,1], regscfac2)
> test2[order(test2[ ,2]), ]
          lifeex...1.   regscfac2
17           Honduras -1.85175707
23       Trinidad(67) -1.10729029
22       Trinidad(62) -0.95989475
3          Madagascar -0.81158820
21             Panama -0.73778272
2             Cameroon -0.72340014
19             Mexico -0.72020002
14          Greenland -0.69123901
16          Guatemala -0.67765804
18             Jamaica -0.51918493
11          CostaRica -0.50500435
25 UnitedStates(NW66) -0.36230253
30           Columbia -0.29543613
4           Mauritius -0.29028454
7      SouthAfrica(C) -0.08065792
12       DominicanRep  0.01111171
8      SouthAfrica(W)  0.06400408
20          Nicaragua  0.08407227
24   UnitedStates(66)  0.16389896
27   UnitedStates(67)  0.16389896
28          Argentina  0.24811968
15            Grenada  0.25241049
10             Canada  0.29564122
26  UnitedStates(W66)  0.40877239
31            Ecuador  0.44246371
5             Reunion  0.47429917
13         ElSalvador  0.65100820
29              Chile  0.75222637
6           Seychelles  0.82902375
1             Algeria  1.90095771
9             Tunisia  3.59177195
> # sort countries by values of third factor scores
> regscfac3 <- model.reg3$scores[,3]
> test3=data.frame(lifeex[ ,1], regscfac3)
> test3[order(test3[ ,2]), ]
          lifeex...1.   regscfac3
2             Cameroon -1.84772224
5             Reunion -1.55046466
```

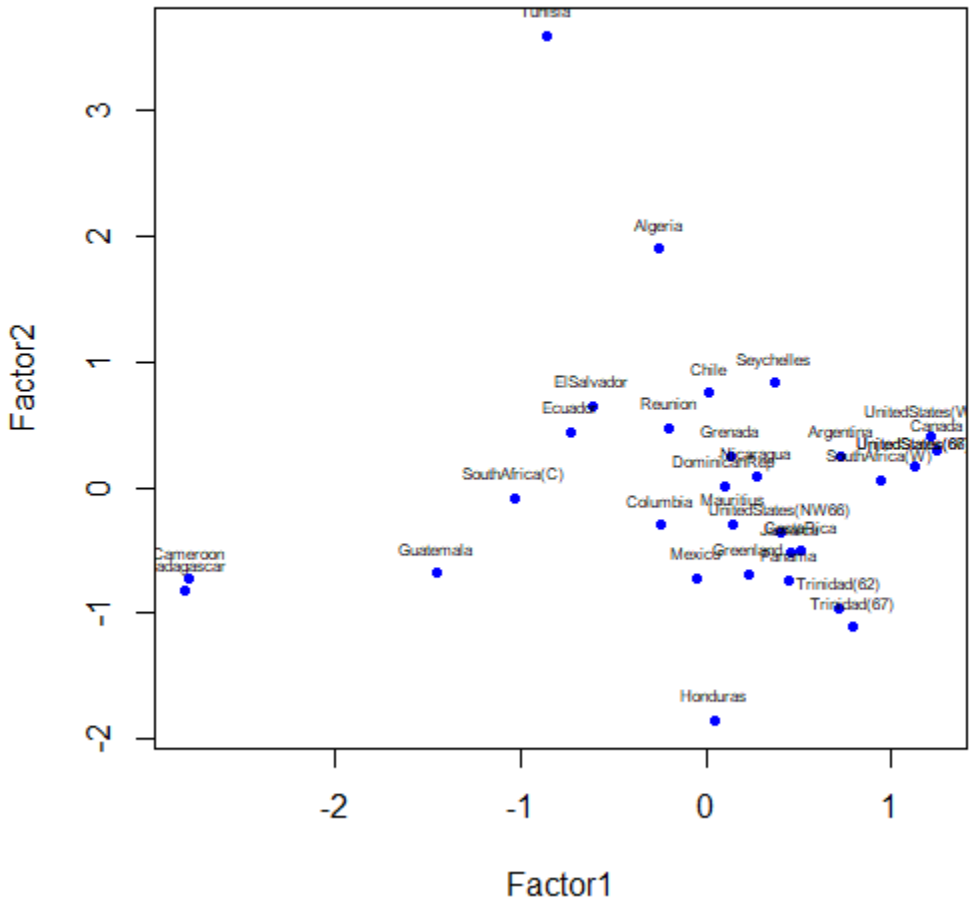
```
8       SouthAfrica(W) -0.91995289
4            Mauritius -0.85862443
25 UnitedStates(NW66) -0.74299137
26  UnitedStates(W66) -0.69225320
24   UnitedStates(66) -0.68177046
27   UnitedStates(67) -0.68177046
7      SouthAfrica(C) -0.65421971
6           Seychelles -0.55214085
23       Trinidad(67) -0.51958264
29              Chile -0.49198911
14          Greenland -0.38558654
9             Tunisia -0.36442148
10             Canada -0.27342781
22       Trinidad(62) -0.21545329
15             Grenada -0.15220645
28           Argentina -0.12817725
3           Madagascar -0.01210318
18             Jamaica  0.08032855
17            Honduras  0.30633182
30            Columbia  0.42919600
19              Mexico  0.44417800
13          ElSalvador  0.48836431
16           Guatemala  0.65911906
11           CostaRica  1.01328707
21              Panama  1.25218728
31             Ecuador  1.59164974
20           Nicaragua  1.70568388
12        DominicanRep  1.83871599
1              Algeria  1.91581631
```

Newborns have a higher estimated life expectancy in Madagascar than the other countries in the list.

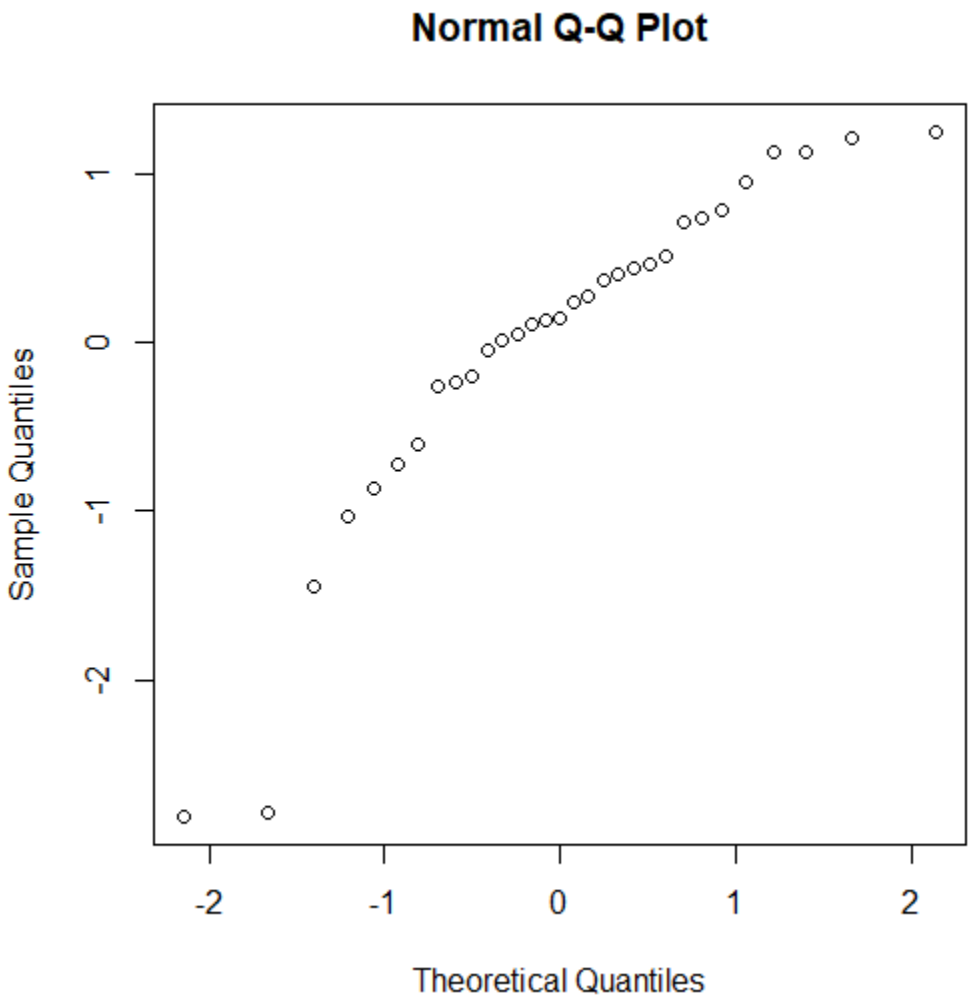
Older women have a higher estimated life expectancy in Honduras than the other countries in the list.

Older men have a higher estimated life expectancy in Cameroon than the other countries in the list.

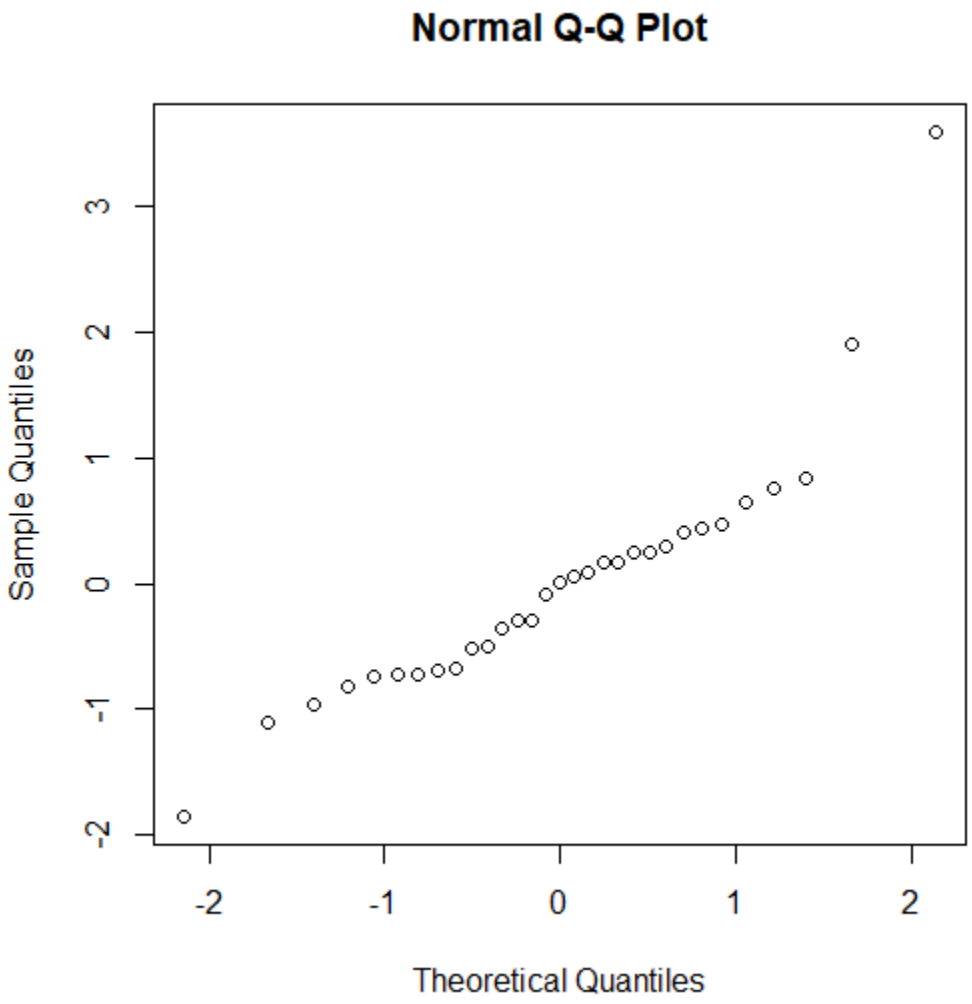
```
# scatterplot of factors 1 and 2 in the regression factor analysis model
> plot(model.reg3$scores[,1:2], pch=20, col="blue")
> text(model.reg3$scores[ ,1], model.reg3$scores[ ,2], lifeex[ ,1], cex=0.5,
pos=3)
```



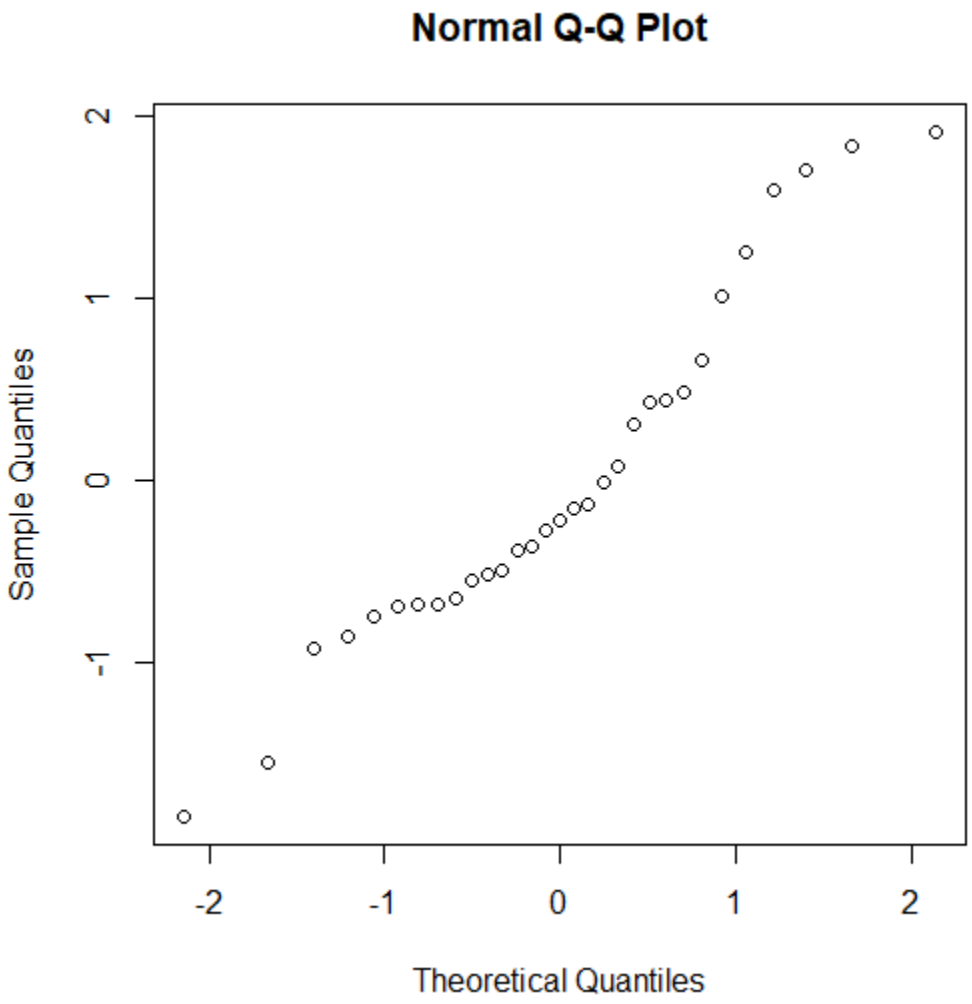

```
# qq plot for 1st factor
> qqnorm(model.reg3$scores[,1])
```




```
# qq plot for 2nd factor
> qqnorm(model.reg3$scores[,2])
```




```
# qq plot for 3rd factor
> qqnorm(model.reg3$scores[,3])
```




The normal assumption holds for the 3-factor model. The points don't deviate much from the straight line which in imagination runs through the middle. For the 2-factor model and 1 factor model, they both seem to deviate from the straight imaginary line. They appear to be following a quadratic shape and don't fit the straight-line requirement of the QQ plots which would make it follow the normal assumption. There do appear to be outliers in the data. For the 3-factor plot there are two outliers in the bottom left, 1 in the top right and bottom left for the 2 factor model. There are two outliers in the bottom right for the 1 factor model.