Lab 1 ***Factor analysis***

Exercise 2
Social mobility in UK

The data are based on information provided by 713 male or female married respondents to a survey carried out in 1949. The variables relate to the respondent, his or his spouse, father, father-in-law, and firstborn son. The file socmob.txt contains the full correlation matrix. The 10 variables are coded as follows:
X1=Husband's father's occupational status
X2=Wife's father's occupational status
X3=Husband's further education
X4=Husband's qualifications
X5=Husband's occupational status
X6=Wife's further education
X7=Wife's qualifications
X8=Firstborn's further education
X9=Firstborn's qualifications
X10=Firstborn's occupational status
Perform first an exploratory factor analysis and then a confirmatory factor analysis on this data set.

1. Load the correlation matrix socmob.txt and explore the correlation matrix in order evaluate if a factor model can be fitted.
   cormat <- read.table("socmob.txt")
   cormat

2. Estimate a factor model with 1,2,3 and 4 factors using the Maximum Likelihood method and select the best model using the Chi-square test. Use the function factanal.
   cormat <- as.matrix(cormat)
   n <- 713
   formula <- "V1 + V2 + V3 + V4 + V5 + V6 + V7 + V8 + V9 + V10"
   f1 <- factanal(formula, factors = 1, covmat = cormat, n.obs = n, rotation = "none")
   f2 <- factanal(formula, factors = 2, covmat = cormat, n.obs = n, rotation = "none")
   f3 <- factanal(formula, factors = 3, covmat = cormat, n.obs = n, rotation = "none")
   f4 <- factanal(formula, factors = 4, covmat = cormat, n.obs = n, rotation = "none")
   names(f1)
   Chisq <- round(c(f1$STATISTIC, f2$STATISTIC, f3$STATISTIC, f4$STATISTIC),

3)
Chisq
df <- c(f1$dof, f2$dof, f3$dof, f4$dof)
pvalues <- round(c(f1$PVAL, f2$PVAL, f3$PVAL, f4$PVAL), 4)
pvalues
f4

3. Compute the communalities and comment them. Which percentage of the variance of the model is explained by the three-factor model?
   comm <- 1 - f3$uniquenesses
   comm
   percVar <- sum(comm)/nrow(cormat)
   percVar

4. Compute the reproduced correlation matrix and the discrepancy between the observed and reproduced correlation.
   repcorr<-loadings(f3)%*%t(loadings(f3))
   round(matcor-repcorr,3)

5. Apply different orthogonal and oblique rotations and interpret the solutions obtained.
   loadings(f3)
   print(f3, cutoff = 0.2)
   library(GPArotation)
   Varimax(loadings(f3))
   quartimax(loadings(f3))
   oblimin(loadings(f3))

6. On the basis of the previous analysis perform a confirmatory factor analysis using the function `cfa` in lavaan.
   library(lavaan)
   socmob.model <- "
   F1 = V8 + V9 + V10
   F2 = V1 + V2 + V5 + V10
   F3 = V3 + V4 + V6 + V7"
   fit <- cfa(socmob.model, sample.cov = cormat, sample.nobs = n, std.lv = TRUE)
   summary(fit, fit.measures = T)