```
##########################################
## solutions for worksheet for module 10 ##
## categorical variable summary stats    ##
##########################################

#########################
###       PART A      ###
#########################

#############
###   1   ###
#############

## first set working directory
setwd("~/Dropbox/AAU materials/datasets/Rosmap")

## double check this
getwd()

## load the data
RM <- read.csv("RM_xsect.csv")

## look at the data
head(RM)


#############
###   2   ###
#############

summary(RM$global_lv)

#############
### 3 (a) ###
#############

?cut
quantile(RM$global_lv, probs=c(0,0.25,0.5,0.75,1), na.rm=TRUE)
RM$global_lv_cat <- cut(RM$global_lv, c(-Inf,-1.83, -0.906, -0.049,
 Inf), right=FALSE, include.lowest = TRUE)
table(RM$global_lv_cat)
```

```
#############
### 3 (b) ###
#############

?factor
RM$global_lv_cat <- factor(RM$global_lv_cat, labels =
 c("Q1","Q2","Q3","Q4"))
table(RM$global_lv_cat)


#############
### 3 (c) ###
#############

quantile(RM$global_lv, probs=c(0,0.25,0.5,0.75,1), na.rm=TRUE)
tapply(RM$global_lv, RM$global_lv_cat, min, na.rm=TRUE)
tapply(RM$global_lv, RM$global_lv_cat, max, na.rm=TRUE)


#############
### 3 (d) ###
#############

tab_global <- table(RM$global_lv_cat)
prop.table(tab_global)
cbind(tab_global, prop.table(tab_global))

#############
### 3 (e) ###
#############


barplot(tab_global, xlab="Global cognative score at last visit",
 ylab="frequency")

###########################
###       PART A        ###
###########################

#############
```

```r
###   1   ###
#############

## load the data
RM <- read.csv("RM_xsect.csv")


#############
###   2   ###
#############

RM$educ_cat <- as.numeric(RM$educ > 12)
table(RM$educ_cat)


#############
### 3 (a) ###
#############

tapply(RM$educ, RM$educ_cat, summary)

#############
### 3 (b) ###
#############

boxplot(educ~educ_cat, data=RM)

#############
###   4   ###
#############

table(RM$educ_cat)
prop.test(1084, 302+1084, p=0.25, correct=FALSE)
```