```
> ### Table 1: knowledge of CV laws in Brazil from survey
>
> library(haven)
>
> ## Clear workspace
>
> rm(list=ls())
>
> ## Set working directory
> setwd("~/Dropbox/Documents/Projects/Active_Projects/Compulsory_Voting_BR/Replication_Files/PSRM/")
>
> ##Opening the data:
>
> load("ReplicationData.RData")
>
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> print_labels(SurveyData$P5A.1)

Labels:
 value      label
     1 Verdadeiro
     2      Falso
    99    Não sei
> knowledge1 <- table(as.numeric(SurveyData$P5A.1 == 1))
> prop.table(knowledge1)*100

        0         1
 6.755725 93.244275
>
> print_labels(SurveyData$P5B.1)

Labels:
 value      label
     1 Verdadeiro
     2      Falso
    99    Não sei
> knowledge3 <- table(as.numeric(SurveyData$P5B.1 == 2))
> prop.table(knowledge3)*100

       0        1
69.08752 30.91248
>
>
> ## Knowledge among young voters (18-24):
>
> DataYoung <- subset(SurveyData, age<25)
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> DataYoung$know1 <- as.numeric(DataYoung$P5A.1 == 1)
> know1Young <- table(DataYoung$know1)
> prop.table(know1Young)*100

        0         1
 6.506365 93.493635
>
> DataYoung$know3 <- as.numeric(DataYoung$P5B.1 == 2)
> know3Young <- table(DataYoung$know3)
> prop.table(know3Young)*100

       0        1
65.94203 34.05797
>
>
> ## Knowledge by education:
>
> ## High school
>
> DataHS <- subset(SurveyData, (P3>=7))
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> DataHS$know1 <- as.numeric(DataHS$P5A.1 == 1)
> know1HS <- table(DataHS$know1)
> prop.table(know1HS)*100

        0         1
 6.456044 93.543956
>
```

```
> DataHS$know3 <- as.numeric(DataHS$P5B.1 == 2)
> know3HS <- table(DataHS$know3)
> prop.table(know3HS)*100

       0        1
67.10467 32.89533
>
>
> ## No high school
>
> DatanoHS <- subset(SurveyData, (P3<7))
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> DatanoHS$know1 <- as.numeric(DatanoHS$P5A.1 == 1)
> know1noHS <- table(DatanoHS$know1)
> prop.table(know1noHS)*100

        0         1
 8.256881 91.743119
>
> DatanoHS$know3 <- as.numeric(DatanoHS$P5B.1 == 2)
> know3noHS <- table(DatanoHS$know3)
> prop.table(know3noHS)*100

       0        1
78.24268 21.75732
>
>
> ## HS vs no HS: binomial proportion test
>
> ## Knowledge item 1:
>
> Edup11 <- know1HS[2]/sum(know1HS)
> Edun11 <- sum(know1HS)
>
> Edup21 <- know1noHS[2]/sum(know1noHS)
> Edun21 <- sum(know1noHS)
>
> Edup1 <- (Edun11 * Edup11 + Edun21 * Edup21)/ (Edun11 + Edun21)
> Eduz1 <- (Edup11 - Edup21) / sqrt(Edup1 * (1-Edup1) * (1/Edun11 + 1/Edun21))
> Eduz1
       1
1.367874
>
> (1 - pnorm(Eduz1, 0, 1))*2
        1
0.1713514
>
> ## Knowledge item 3:
>
> Edup13 <- know3HS[2]/sum(know3HS)
> Edun13 <- sum(know3HS)
>
> Edup23 <- know3noHS[2]/sum(know3noHS)
> Edun23 <- sum(know3noHS)
>
> Edup3 <- (Edun13 * Edup13 + Edun23 * Edup23)/ (Edun13 + Edun23)
> Eduz3 <- (Edup13 - Edup23) / sqrt(Edup3 * (1-Edup3) * (1/Edun13 + 1/Edun23))
> Eduz3
       1
4.777315
>
> (1 - pnorm(Eduz3, 0, 1))*2
           1
1.776512e-06
>
>
>
> ## Gender differences: Online Appendix section S5.1
>
> print_labels(SurveyData$sex)

Labels:
 value  label
     1  Homem
     2 Mulher
>
> ## Knowledge by gender among young voters:
>
> ## Women
```

```
>
> DataWomen <- subset(SurveyData, (sex == 2 & age<25))
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> DataWomen$know1 <- as.numeric(DataWomen$P5A.1 == 1)
> know1Women <- table(DataWomen$know1)
> prop.table(know1Women)*100

        0         1
 5.299539 94.700461
>
> DataWomen$know3 <- as.numeric(DataWomen$P5B.1 == 2)
> know3Women <- table(DataWomen$know3)
> prop.table(know3Women)*100

       0        1
61.95652 38.04348
>
>
> ## Men
>
> DataMen <- subset(SurveyData, (sex == 1 & age<25))
>
> ##Frequencies of knowledge questions: 1 correct answer, 0 otherwise
>
> DataMen$know1 <- as.numeric(DataMen$P5A.1 == 1)
> know1Men <- table(DataMen$know1)
> prop.table(know1Men)*100

        0         1
 8.424908 91.575092
>
> DataMen$know3 <- as.numeric(DataMen$P5B.1 == 2)
> know3Men <- table(DataMen$know3)
> prop.table(know3Men)*100

       0        1
70.49689 29.50311
>
>
> ## Women vs Men: binomial proportion test
>
> ## Knowledge item 1:
>
> Edup11 <- know1Women[2]/sum(know1Women)
> Edun11 <- sum(know1Women)
>
> Edup21 <- know1Men[2]/sum(know1Men)
> Edun21 <- sum(know1Men)
>
> Edup1 <- (Edun11 * Edup11 + Edun21 * Edup21)/ (Edun11 + Edun21)
> Eduz1 <- (Edup11 - Edup21) / sqrt(Edup1 * (1-Edup1) * (1/Edun11 + 1/Edun21))
> Eduz1
       1
1.64043
>
> (1 - pnorm(Eduz1, 0, 1))*2
         1
0.1009159
>
> ## Knowledge item 3:
>
> Edup13 <- know3Women[2]/sum(know3Women)
> Edun13 <- sum(know3Women)
>
> Edup23 <- know3Men[2]/sum(know3Men)
> Edun23 <- sum(know3Men)
>
> Edup3 <- (Edun13 * Edup13 + Edun23 * Edup23)/ (Edun13 + Edun23)
> Eduz3 <- (Edup13 - Edup23) / sqrt(Edup3 * (1-Edup3) * (1/Edun13 + 1/Edun23))
> Eduz3
       1
2.361639
>
> (1 - pnorm(Eduz3, 0, 1))*2
          1
0.01819433
>
>
>
```