

“Convert Numeric Vectors to Coded Factors”

Chenoa Schatzki-McClain

April 4, 2018

```
library(stringr)
```

```
## Warning: package 'stringr' was built under R version 3.4.3
```

```
# Load IFLS4 <- numeric vectors
```

```
IFLS4 <- read.csv("C:/Users/cheno/Desktop/IFLS_all/DATA/R datasets/IFLS4.csv")
```

```
# Load IFLS5 <- labeled factors
```

```
IFLS5 <- read.csv("C:/Users/cheno/Desktop/IFLS_all/DATA/R datasets/IFLS5.csv")
```

```
#list of vars for which type == class in IFLS5
```

```
vars <- names(IFLS5)[sapply(IFLS5[,names(IFLS5)], class) == "factor"]
```

```
relevel_IFLS5 <- function(var){
```

```
  lev <- levels(IFLS5[,var])
```

```
  num <- as.numeric(str_extract(levels(IFLS5[,var]), "\\d+(?=:)"))
```

```
  num <- sort(num)
```

```
  lev2 <- vector(length = length(lev))
```

```
  for (i in 1:length(lev)){
```

```
    lev2[i] <- lev[as.numeric(str_extract(levels(IFLS5[,var]), "\\d+(?=:)")) == num[i]]
```

```
  }
```

```
  return(lev2)
```

```
}
```

```
for (i in 1:length(vars)){
```

```
  levels(IFLS5[,vars[i]]) <- relevel_IFLS5(vars[i])
```

```
}
```

```
convert <- function(var){
```

```
  temp <- as.factor(IFLS4[,var])
```

```
#check that all levels are present and the same in both datasets >> could still be issues if coding i  
if(length(levels(temp)) == length(levels(IFLS5[,var]))){
```

```
  if(sum(levels(temp) == str_extract(levels(IFLS5[,var]), "\\d+(?=:)")) == length(levels(temp))){
```

```
    levels(temp) <- levels(IFLS5[,var])}
```

```
  return(temp)
```

```
}
```

```
for (i in 1:length(vars)){
```

```
  IFLS4[,i] <- convert(i)
```

```
}
```

```
#sum(levels(as.factor(IFLS4$dl06x_H)) == substr(levels(IFLS5$dl06x_H), 1, 1)) == length(levels(as.factor
```

```
#length(levels(as.factor(IFLS4$dl06x_H))) == length(levels(IFLS5$dl06x_H))
```

```
#issue here
```

```
#sum(levels(as.factor(IFLS4$tk46dm)) == substr(levels(IFLS5$tk46dm), 1, 1)) == length(levels(as.factor(I
#length(levels(as.factor(IFLS4$tk46dm))) == length(levels(IFLS5$tk46dm))
#length(levels(IFLS4$tk01a)) == length(levels(IFLS5[, "tk01a"])))
```

```
levels(IFLS4$tk01a)
```

```
## [1] "1" "3" "9"
```

```
levels(IFLS5$tk01a)
```

```
## [1] "1:Yes" "3:No"
```

```
show_levels <- function(var){
  temp <- IFLS5[,var]
  return(levels(temp))
}
```

```
sapply(vars, show_levels)
```

```
## $sex
## [1] "3:Female"
##
## $rspndnt
## [1] "1:Head of household" "2:Spouse of head" "3:Other HH member"
##
## $marstat
## [1] "2:Married"
##
## $panel
## [1] "1:Panel" "3:New"
##
## $tk01
## [1] "1:Working/trying to get work/helping to earn income"
## [2] "3:Attending school"
## [3] "4:Housekeeping"
## [4] "5:Retired"
## [5] "7:Sick/disable"
## [6] "95:Other"
##
## $tk01a
## [1] "1:Yes" "3:No"
##
## $tk03
## [1] "1:Yes" "3:No"
##
## $tk04
## [1] "1:Yes"
##
## $tk24a
## [1] "1:Self-employed"
## [2] "2:Self employed with unpaid family/temporary worker"
```

```

## [3] "3:Self-employed with employees/permanent workers"
## [4] "4:Government worker"
## [5] "5:Private worker"
## [6] "6:Unpaid family worker"
## [7] "7:Casual worker in agriculture"
## [8] "8:Casual worker not in agriculture"
## [9] "9:Missing"
##
## $tk46a
## [1] "1:Yes" "3:No"
##
## $tk46b
## [1] "1:Yes" "3:No"
##
## $tk46dm
## [1] "1:January"      "2:February"      "3:March"          "4:April"
## [5] "5:May"          "6:June"          "7:July"           "8:August"
## [9] "9:September"    "10:October"      "11:November"      "12:December"
## [13] "98:Don't Know"
##
## $tk46m
## [1] "1:Fire by company because business was closed down"
## [2] "2:Fired for other reason"
## [3] "3:Wage/salary was too low"
## [4] "4:Not conducive working environment"
## [5] "5:Refused being relocated"
## [6] "6:Prolonged sickness"
## [7] "7:Marriage"
## [8] "8:Childbirth"
## [9] "9:Other family reason"
## [10] "95:Other"
##
## $dl06x
## [1] "1:Yes" "3:No"
##
## $dl06
## [1] "2:Elementary school"
## [2] "3:Junior high general"
## [3] "4:Junior high vocational"
## [4] "5:Senior high general"
## [5] "6:Senior high vocational"
## [6] "11:Adult education A"
## [7] "12:Adult education B"
## [8] "13:Open university"
## [9] "14:Islamic School (pesantren)"
## [10] "15:Adult education C"
## [11] "17:School for Disabled"
## [12] "60:College (D1,D2,D3)"
## [13] "61:University S1"
## [14] "62:University S2"
## [15] "63:University S3"
## [16] "72:Islamic Elementary School (Madrasah Ibtidaiyah)"
## [17] "73:Islamic Junior/High School (Madrasah Tsanawiyah)"
## [18] "74:Islamic Senior/High School (Madrasah Tsanawiyah)"

```

```

## [19] "90:Kindergarten"
## [20] "95:Other"
## [21] "98:Don't Know"
## [22] "99:MISSING"
##
## $ar13
## [1] "1:Unmarried" "2:Married"
##
## $tk24a_H
## [1] "1:Self-employed"
## [2] "2:Self employed with unpaid family/temporary worker"
## [3] "3:Self-employed with employees/permanent workers"
## [4] "4:Government worker"
## [5] "5:Private worker"
## [6] "6:Unpaid family worker"
## [7] "7:Casual worker in agriculture"
## [8] "8:Casual worker not in agriculture"
##
## $tk46a_H
## [1] "1:Yes" "3:No"
##
## $tk46b_H
## [1] "1:Yes" "3:No"
##
## $tk46m_H
## [1] "1:Fire by company because business was closed down"
## [2] "2:Fired for other reason"
## [3] "3:Wage/salary was too low"
## [4] "4:Not conducive working environment"
## [5] "5:Refused being relocated"
## [6] "6:Prolonged sickness"
## [7] "7:Marriage"
## [8] "8:Childbirth"
## [9] "9:Other family reason"
## [10] "95:Other"
##
## $tk01_H
## [1] "1:Working/trying to get work/helping to earn income"
## [2] "2:Job searching"
## [3] "3:Attending school"
## [4] "4:Housekeeping"
## [5] "5:Retired"
## [6] "7:Sick/disable"
## [7] "95:Other"
##
## $tk01a_H
## [1] "1:Yes" "3:No"
##
## $tk03_H
## [1] "1:Yes" "3:No"
##
## $tk04_H
## [1] "1:Yes"
##

```

```

## $dl06_H
## [1] "2:Elementary school"
## [2] "3:Junior high general"
## [3] "4:Junior high vocational"
## [4] "5:Senior high general"
## [5] "6:Senior high vocational"
## [6] "11:Adult education A"
## [7] "12:Adult education B"
## [8] "13:Open university"
## [9] "14:Islamic School (pesantren)"
## [10] "15:Adult education C"
## [11] "17:School for Disabled"
## [12] "60:College (D1,D2,D3)"
## [13] "61:University S1"
## [14] "62:University S2"
## [15] "63:University S3"
## [16] "72:Islamic Elementary School (Madrasah Ibtidaiyah)"
## [17] "73:Islamic Junior/High School (Madrasah Tsanawiyah)"
## [18] "74:Islamic Senior/High School (Madrasah Tsanawiyah)"
## [19] "95:Other"
## [20] "98:Don't Know"
##
## $dl06x_H
## [1] "1:Yes" "3:No"
##
## $sc05
## [1] "1:Urban" "2:Rural"

```