Data Summaries

Chenoa Schatzki-McClain

April 10, 2018

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
rm(list=ls())
```

```
#load data
clean <- read.csv("C:/Users/cheno/Desktop/IFLS_all/DATA/R datasets/clean_4_10.csv")

#filter to only include urban households
urban <- clean %>%
        filter(sc05 == "1:Urban")

#filter to show only husbands from urban HH that lost jobs
urban_jl <- clean %>%
        filter (sc05 == "1:Urban" & job_loss_H == "1:Yes")

#filter to urban HH w/ employed husband
emp_wives <- urban %>%
        filter(employed == "1:Employed")

#filter to urban HH w/ employed husband
emp_husbands <- urban %>%
        filter(employed_H == "1:Employed")
```

```
#function to produce table summaries

sum_tab <- function(var){

  prop.table(table(var))

}
```

```
#wives' ages
summary(urban$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   15.00   31.00   38.00   39.46   47.00   95.00
```

```
#ages of employed wives
summary(emp_wives$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   16.00   32.00   39.00   39.89   47.00   82.00
```

```
#husbands' ages
summary(urban$age_H)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   16.00   35.00   42.00   43.97   52.00   96.00     815
```

```
#ages of employed  husbands
summary(emp_wives$age_H)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   19.00   36.00   43.00   44.26   52.00   88.00     430
```

```
#wives'educational attainment
summary(urban$dl06)
```

```
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##            335           2734           1813           2811           1189
##           NA's
##            292
```

```
sum_tab(urban$dl06)
```

```
## var
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##     0.03771673     0.30781356     0.20412069     0.31648277     0.13386625
```

```
#educational attainment of employed wives
summary(emp_wives$dl06)
```

```
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##            201           1599            969           1467            887
##           NA's
##              1
```

```
sum_tab(emp_wives$dl06)
```

```
## var
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##     0.03923482     0.31212180     0.18914698     0.28635565     0.17314074
```

```
#husbands' educational attainment
summary(urban$dl06_H)
```

```
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##            150           2323           1321           2873           1270
##           NA's
##           1237
```

```
sum_tab(urban$dl06_H)
```

```
## var
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##     0.01889883     0.29267985     0.16643568     0.36197556     0.16001008
```

```
#educational attainment of employed husbands
summary(emp_husbands$dl06_H)
```

```
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
```

```
##            129           2087           1204           2685           1159
##           NA's
##              2
```

```
sum_tab(emp_husbands$dl06_H)
```

```
## var
## 1:no schooling   2:elementary      3:juniorH      4:seniorH       5:higher
##     0.01775881     0.28730727     0.16574890     0.36963106     0.15955396
```

```
#wives employment status
summary(urban$employed)
```

```
##   1:Employed 2:Unemployed
##         5124         4050
```

```
sum_tab(urban$employed)
```

```
## var
##   1:Employed 2:Unemployed
##     0.558535     0.441465
```

```
#husbands' employment status
summary(urban$employed_H)
```

```
##   1:Employed 2:Unemployed
##         7266         1908
```

```
sum_tab(urban$employed_H)
```

```
## var
##   1:Employed 2:Unemployed
##    0.7920209    0.2079791
```

```
#sector of employment, working wives
table(emp_wives$tk24a)
```

```
##
##                                      1:Self-employed
##                                                  974
## 2:Self employed with  unpaid family/temporary worker
##                                                  922
##     3:Self-employed with employees/permanent workers
##                                                   98
##                                    4:Government worker
##                                                  461
##                                       5:Private worker
##                                                 1565
##                                  6:Unpaid family worker
##                                                  706
##                         7:Casual worker in agriculture
##                                                  101
##                     8:Casual worker not in agriculture
##                                                  277
```

```
sum_tab(emp_wives$tk24a)
```

```
## var
##                                      1:Self-employed
```

```
##                                          0.19083072
## 2:Self employed with  unpaid family/temporary worker
##                                          0.18064263
##     3:Self-employed with employees/permanent workers
##                                          0.01920063
##                                  4:Government worker
##                                          0.09032132
##                                     5:Private worker
##                                          0.30662226
##                                6:Unpaid family worker
##                                          0.13832288
##                       7:Casual worker in agriculture
##                                          0.01978840
##                    8:Casual worker not in agriculture
##                                          0.05427116
```

```
#sector of employment, working husbands
table(emp_husbands$tk24a)
```

```
##
##                                       1:Self-employed
##                                                   810
## 2:Self employed with  unpaid family/temporary worker
##                                                   817
##     3:Self-employed with employees/permanent workers
##                                                    73
##                                     4:Government worker
##                                                   377
##                                        5:Private worker
##                                                  1305
##                                   6:Unpaid family worker
##                                                   703
##                          7:Casual worker in agriculture
##                                                    89
##                       8:Casual worker not in agriculture
##                                                   227
```

```
sum_tab(emp_husbands$tk24a)
```

```
## var
##                                       1:Self-employed
##                                          0.18404908
## 2:Self employed with  unpaid family/temporary worker
##                                          0.18563963
##     3:Self-employed with employees/permanent workers
##                                          0.01658714
##                                  4:Government worker
##                                          0.08566235
##                                     5:Private worker
##                                          0.29652352
##                                6:Unpaid family worker
##                                          0.15973642
##                       7:Casual worker in agriculture
##                                          0.02022268
##                    8:Casual worker not in agriculture
```

```
##                                         0.05157919
```

```
#wage wives
summary(emp_wives$tk25a1)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
##        0   300000   720000  1231576  1650000 23000000     2732
```

```
#wage husbands
summary(emp_husbands$tk25a1_H)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
##        0   700000  1300000  1847224  2275250 76000000     2882
```

```
#number of dependents
summary(urban$dependents)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   1.000   1.301   2.000   8.000     611
```

```
#working dependents
summary(urban$working_dependents)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.0000  0.0000  0.0078  0.0000  2.0000     611
```

```
#number of other HH members
summary(urban$other_HHM)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   4.000   5.066   7.000  23.000     611
```

```
#other working HH members
summary(urban$other_working)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   2.000   2.624   3.000  17.000    2611
```

```
#region >> need to locate codes, remove region 15?
table(urban$sc_code)
```

```
##
##   12   13   14   15   16   18   19   21   31   32   33   34   35   36   51
##  570  438   61    1  250  184   85   18 1040 1563  927  688 1274  325  514
##   52   62   63   64   73   76
##  470   13  329   22  378   24
```

```
#INFO on HUSBANDS WHO LOSE JOB

#husbands' job losses
table(urban$job_loss)
```

```
##
## 1:Yes  2:No
##   452  5144
```

```
sum_tab(urban$job_loss)
```

```
## var
##      1:Yes       2:No
## 0.08077198 0.91922802
```

```r
#age of husbands who lost jobs
summary(urban_jl$age_H)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   22.00   31.00   36.00   38.03   43.00   75.00
```

```r
#current employment status of husbands who lost jobs
summary(urban_jl$employed_H)
```

```
##   1:Employed 2:Unemployed
##          868           33
```

```r
sum_tab(urban_jl$employed_H)
```

```
## var
##   1:Employed 2:Unemployed
##   0.96337403   0.03662597
```

```r
#current sector of employment, husbands who lost jobs
summary(urban_jl$tk24a_H)
```

```
##                                      1:Self-employed
##                                                  115
## 2:Self employed with  unpaid family/temporary worker
##                                                   80
##     3:Self-employed with employees/permanent workers
##                                                   17
##                                   4:Government worker
##                                                   23
##                                      5:Private worker
##                                                  540
##                                 6:Unpaid family worker
##                                                   23
##                         7:Casual worker in agriculture
##                                                    6
##                    8:Casual worker not in agriculture
##                                                   63
##                                                NA's
##                                                   34
```

```r
sum_tab(urban_jl$tk24a_H)
```

```
## var
##                                      1:Self-employed
##                                          0.132641292
## 2:Self employed with  unpaid family/temporary worker
##                                          0.092272203
##     3:Self-employed with employees/permanent workers
##                                          0.019607843
##                                   4:Government worker
##                                          0.026528258
##                                      5:Private worker
##                                          0.622837370
##                                 6:Unpaid family worker
##                                          0.026528258
##                         7:Casual worker in agriculture
##                                          0.006920415
```

```
##                   8:Casual worker not in agriculture
##                                          0.072664360
```

```
#for husbands who lost jobs before second period and were employed in 1st period, job type in first per

#husbands who lost jobs before IFLS5
jl_IFLS5 <- urban_jl %>%
              filter(year == 2014)

#husbands who lost jobs before IFLS4
jl_IFLS4 <- urban_jl %>%
              filter(year == 2007)


#husbands employed in IFLS4 who lost jobs in IFLS5 (IFLS4 data)
emp_H_IFLS4_jl <- emp_husbands %>%
                filter(year == 2007, pidlink %in% jl_IFLS5$pidlink)

#husbands employed in IFLS4 who lost jobs in IFLS5 (IFLS4 data)
jl_IFLS5_2 <- jl_IFLS5 %>%
              filter(pidlink %in% emp_H_IFLS4_jl$pidlink)


#industry that husbands worked in prior to job loss >> disproportionately private sector
summary(emp_H_IFLS4_jl$tk24a_H)
```

```
##                                      1:Self-employed
##                                                   28
## 2:Self employed with  unpaid family/temporary worker
##                                                   12
##     3:Self-employed with employees/permanent workers
##                                                    2
##                                    4:Government worker
##                                                    4
##                                       5:Private worker
##                                                  229
##                                  6:Unpaid family worker
##                                                    4
##                         7:Casual worker in agriculture
##                                                    8
##                   8:Casual worker not in agriculture
##                                                   26
```

```
sum_tab(emp_H_IFLS4_jl$tk24a_H)
```

```
## var
##                                      1:Self-employed
##                                          0.089456869
## 2:Self employed with  unpaid family/temporary worker
##                                          0.038338658
##     3:Self-employed with employees/permanent workers
##                                          0.006389776
##                                    4:Government worker
##                                          0.012779553
##                                       5:Private worker
```

```
##                                       0.731629393
##                                6:Unpaid family worker
##                                       0.012779553
##                          7:Casual worker in agriculture
##                                       0.025559105
##                     8:Casual worker not in agriculture
##                                       0.083067093
```

```
#industry that husbands worked in after job loss >> slight decrease in private sector
summary(jl_IFLS5_2$tk24a_H)
```

```
##                                      1:Self-employed
##                                                   39
## 2:Self employed with  unpaid family/temporary worker
##                                                   32
##     3:Self-employed with employees/permanent workers
##                                                    8
##                                    4:Government worker
##                                                    4
##                                       5:Private worker
##                                                  200
##                                6:Unpaid family worker
##                                                    6
##                          7:Casual worker in agriculture
##                                                    1
##                     8:Casual worker not in agriculture
##                                                   23
```

```
sum_tab(jl_IFLS5_2$tk24a_H)
```

```
## var
##                                      1:Self-employed
##                                         0.124600639
## 2:Self employed with  unpaid family/temporary worker
##                                         0.102236422
##     3:Self-employed with employees/permanent workers
##                                         0.025559105
##                                    4:Government worker
##                                         0.012779553
##                                       5:Private worker
##                                         0.638977636
##                                6:Unpaid family worker
##                                         0.019169329
##                          7:Casual worker in agriculture
##                                         0.003194888
##                     8:Casual worker not in agriculture
##                                         0.073482428
```

```
#husbands who lost jobs in  both periods?  >> 78
nrow(jl_IFLS5 %>%
       filter(pidlink %in% jl_IFLS4$pidlink))
```

```
## [1] 78
```

```
#average number of job losses  >> 39 is unreasonable, will need to revisit this
summary(urban_jl$tk46c_H)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   1.000   1.000   1.516   2.000  39.000       3
```