Highest Degree and Political Party Affiliation

mickfenneck

22 settembre 2014

Introduction:

This analysis is thought for analyzing a General Social Survey in order to find if exists a correlation between the Political Party Affiliation of a subject and its highest degree.

Data:

This extract of the General Social Survey (GSS) Cumulative File 1972-2012 provides a sample of selected indicators in the GSS with the goal of providing a convenient data resource for students learning statistical reasoning using the R language. Unlike the full General Social Survey Cumulative File, we have removed missing values from the responses and created factor variables when appropriate to facilitate analysis using R. Our hope is that this will allow students to focus on statistical concepts without having to (initially) be concerned about some of the data management and interpretation issues associated with missing data and factor variables in R. Other than the two modifications mentioned above, all data and coding come from the original dataset. Students and researchers seeking to conduct research or explore the full codebook behind the full General Social Survey Cumulative File are urged to consult the original dataset at the citation that follows:

Smith, Tom W., Michael Hout, and Peter V. Marsden. General Social Survey, 1972-2012 [Cumulative File]. ICPSR34802-v1. Storrs, CT: Roper Center for Public Opinion Research, University of Connecticut /Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors], 2013-09-11. doi:10.3886/ICPSR34802.v1

Exploratory data analysis:

The basic structure of the dataset is

```
names(gss)
```

```
##   [1] "caseid"   "year"     "age"      "sex"      "race"     "hispanic"
##   [7] "uscitzn"  "educ"     "paeduc"   "maeduc"   "speduc"   "degree"
##  [13] "vetyears" "sei"      "wrkstat"  "wrkslf"   "marital"  "spwrksta"
##  [19] "sibs"     "childs"   "agekdbrn" "incom16"  "born"     "parborn"
##  [25] "granborn" "income06" "coninc"   "region"   "partyid"  "polviews"
##  [31] "relig"    "attend"   "natspac"  "natenvir" "natheal"  "natcity"
##  [37] "natcrime" "natdrug"  "nateduc"  "natrace"  "natarms"  "nataid"
##  [43] "natfare"  "natroad"  "natsoc"   "natmass"  "natpark"  "confinan"
##  [49] "conbus"   "conclerg" "coneduc"  "confed"   "conlabor" "conpress"
##  [55] "conmedic" "contv"    "conjudge" "consci"   "conlegis" "conarmy"
##  [61] "joblose"  "jobfind"  "satjob"   "richwork" "jobinc"   "jobsec"
##  [67] "jobhour"  "jobpromo" "jobmeans" "class"    "rank"     "satfin"
##  [73] "finalter" "finrela"  "unemp"    "govaid"   "getaid"   "union"
##  [79] "getahead" "parsol"   "kidssol"  "abdefect" "abnomore" "abhlth"
##  [85] "abpoor"   "abrape"   "absingle" "abany"    "pillok"   "sexeduc"
##  [91] "divlaw"   "premarsx" "teensex"  "xmarsex"  "homosex"  "suicide1"
##  [97] "suicide2" "suicide3" "suicide4" "fear"     "owngun"   "pistol"
```

```
## [103] "shotgun"  "rifle"    "news"     "tvhours"  "racdif1"  "racdif2"
## [109] "racdif3"  "racdif4"  "helppoor" "helpnot"  "helpsick" "helpblk"
```

As can be seen the dataset provides many different variable but the only one in which we are interested are "degree"" and "partyid". We can subset that variables with

```
degree <- gss$degree
partyid <- gss$partyid

tbdata <- table(degree, partyid)

summary(tbdata)
```

```
## Number of cases in table: 55754
## Number of factors: 2
## Test for independence of all factors:
##  Chisq = 1473, df = 28, p-value = 5e-293
```

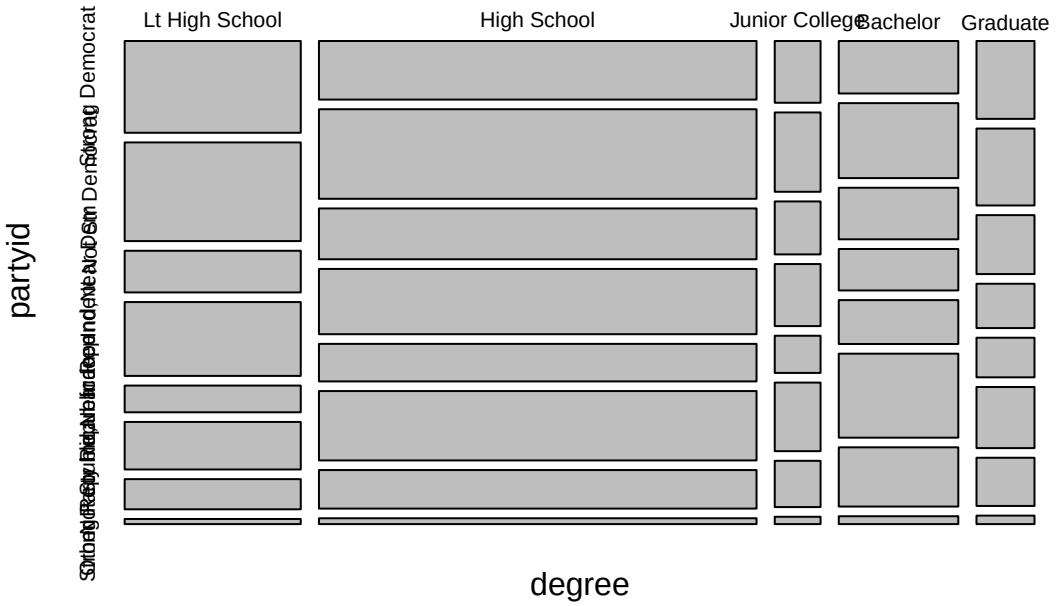
We have 55754 numbers of cases and 2 factor categorical values.

We can see the entire table and plot it.

```
tbdata
```

```
##                 partyid
## degree           Strong Democrat Not Str Democrat Ind,Near Dem Independent
##   Lt High School            2591             2786         1176        2085
##   High School               4110             6281         3565        4579
##   Junior College             454              583          389         456
##   Bachelor                  1006             1436          988         795
##   Graduate                   724              715          548         412
##                 partyid
## degree           Ind,Near Rep Not Str Republican Strong Republican
##   Lt High School          754               1342               852
##   High School            2633               4865              2702
##   Junior College          273                504               340
##   Bachelor                841               1610              1135
##   Graduate                368                568               448
##                 partyid
## degree           Other Party
##   Lt High School         145
##   High School            411
##   Junior College          54
##   Bachelor               151
##   Graduate                79
```

```
plot(tbdata, main="Party Affiliation")
```



Inference:

Insert inference section here...

Conclusion:

Insert conclusion here...