Total Family Income vs Abpoor (Low income - cant afford more children)

Introduction:

1) Research Question: Is there a relationship between total family income (income06) and likelihood of having an abortion?

2) This question can give an insight that, if we consider any family situation whose total family income is less than $1000 and if the mother is pregnant, then the family cannot afford to grow the child which will likely lead to aborting the child. So, I care that if there is any such kind of relationship seen, then something has to be done inorder to reduce the abortion rate.

3) A pregnant mother should not take the decision of abortion due to low family income which can be unaccepatble reason for abortion. This case has to be seriously taken care if we bring out a positive inference in this research question.

Data:

1) Data Collection: The data consist of GSS Cumulative File from 1972 - 2012. General Social Survey (GSS) has been monitoring societal change and studying the growing complexity of American society. The GSS aims to gather data on contemporary American society in order to monitor and explain trends and constants in attitudes, behaviors, and attributes; to examine the structure and functioning of society in general as well as the role played by relevant subgroups; to compare the United States to other societies in order to place American society in comparative perspective and develop cross-national models of human society; and to make high-quality data easily accessible to scholars, students, policy makers, and others, with minimal cost and waiting.

GSS questions cover a diverse range of issues including national spending priorities, marijuana use, crime and punishment, race relations, quality of life, confidence in institutions, and sexual behavior.

2) The cases are the individuals in USA.

3) The two variables used in the study are “incom06” and “abpoor” and both are categorical variables.

4) The type of study used here is “observational study” as the data collected does not directly interfere with how the data arise. As already stated, the data are collected randomly and it has nothing to do how data looks like.

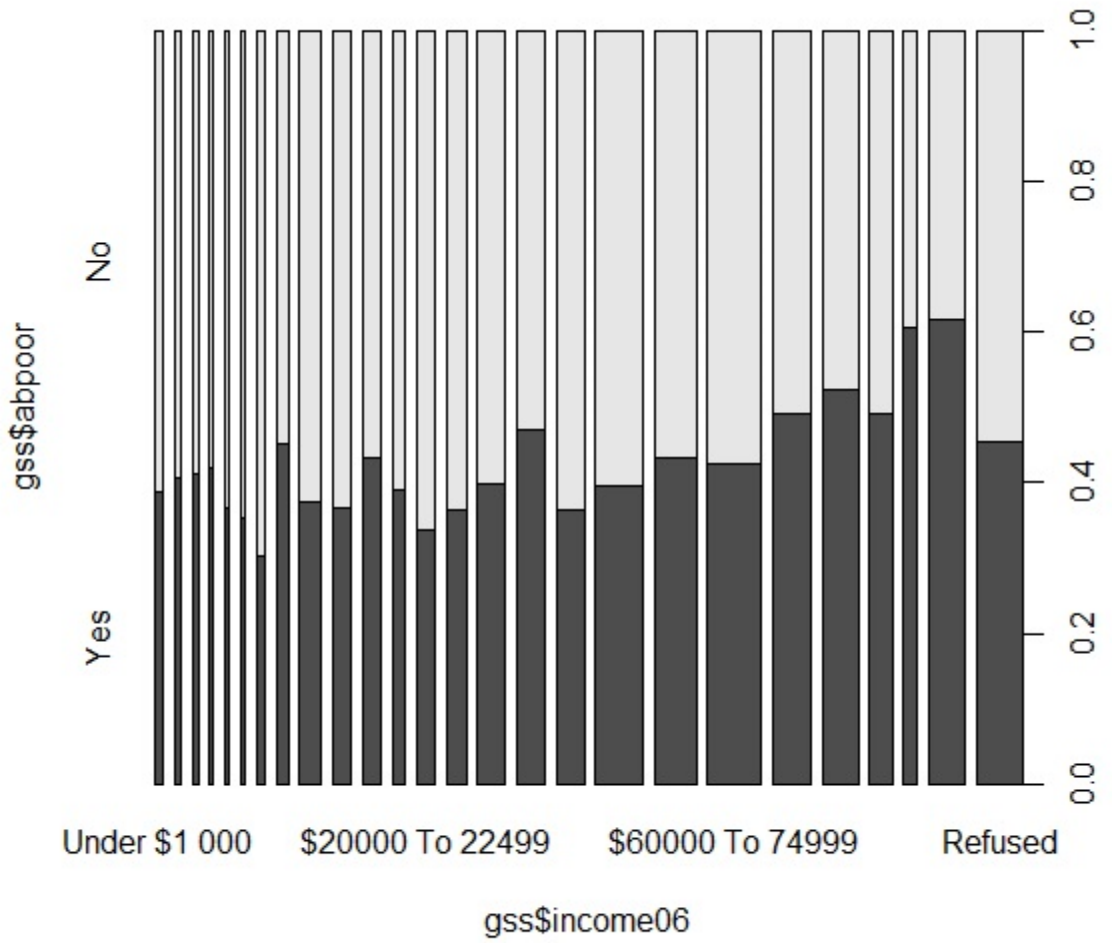
5) The population of interest is all the people in US whose income is under $1000. We can generalize the findings from this analysis and generalize to the whole population. Once we find there is positive relation in abpoor and income under $1000, we can generalize to other income group and find whether it is alternative hypothesis or null hypothesis.

6) The data cannot be used to produce causal link as this is an observational study. In an observational study, we can establish association between the explanatory and response variables.

Exploratory data analysis:

-> Let us plot the abpoor variable with the income06 variable.

```
plot(gss$abpoor ~ gss$income06)
```



-> From the above plot we get a picture of how many people in each income category answered “yes” or “no” in abpoor field.

-> Let us see how many of the people in each category surveyed “yes” or “no” in abpoor for each income group. The below code help us to get the count.

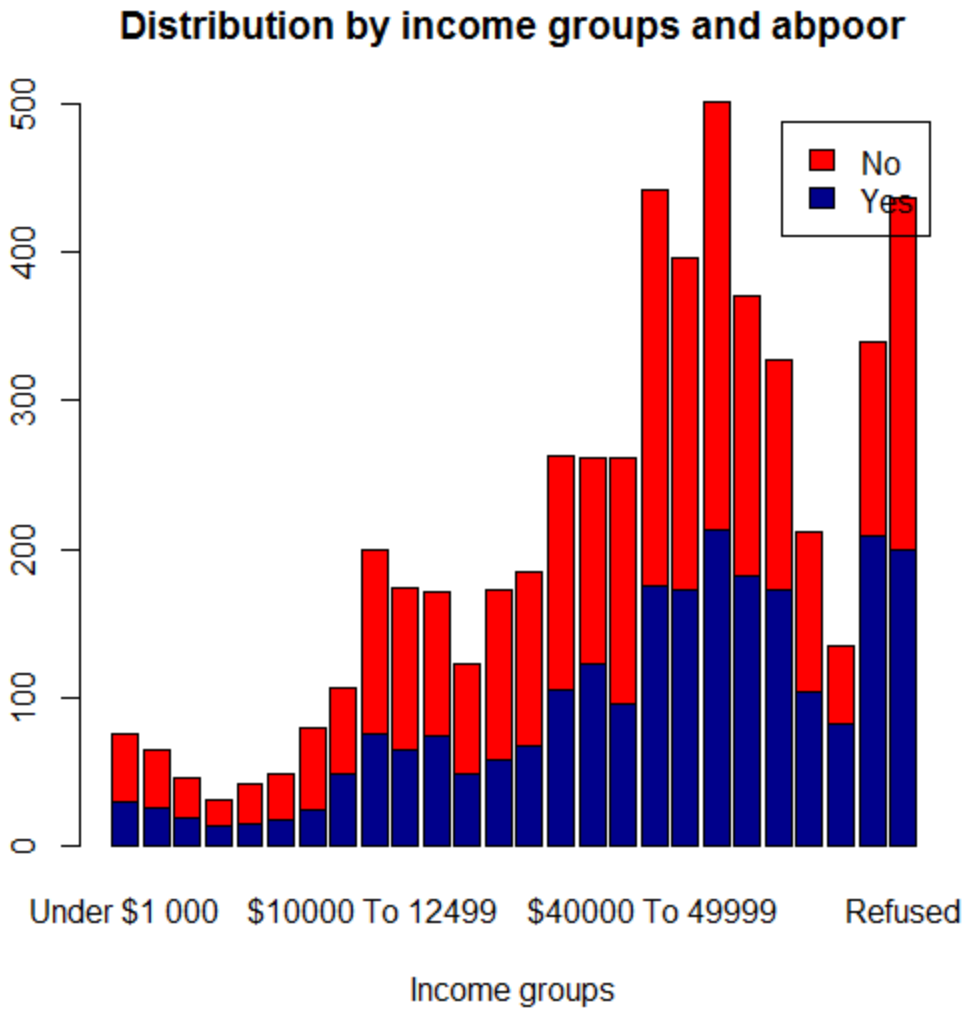
```
mytable <- table(gss$abpoor, gss$income06)
mytable
##
##       Under $1 000 $1 000 To 2 999 $3 000 To 3 999 $4 000 To 4 999
##   Yes           29              26              19              13
##   No            46              38              27              18
##
##       $5 000 To 5 999 $6 000 To 6 999 $7 000 To 7 999 $8 000 To 9 999
##   Yes              15              17              24              48
##   No               26              31              55              58
##
##       $10000 To 12499 $12500 To 14999 $15000 To 17499 $17500 To 19999
##   Yes              75              64              74              48
##   No              125             110              97              75
##
##       $20000 To 22499 $22500 To 24999 $25000 To 29999 $30000 To 34999
```

```
##   Yes               58               67              105              123
##   No               114              117              158              138
##
##       $35000 To 39999 $40000 To 49999 $50000 To 59999 $60000 To 74999
##   Yes               95              175              172              213
##   No               167              267              224              288
##
##       $75000 To $89999 $90000 To $109999 $110000 To $129999
##   Yes               182               172                104
##   No                189               156                108
##
##       $130000 To $149999 $150000 Or Over Refused
##   Yes                 82              209     199
##   No                  53              130     238
```

-> From the above table, we can see that there are about 29 people who survyed “yes” and 49 people surveyed “no”. Likewise we can see the count for other income groups as well. But we cannot come to any conclusion on the relationship without doing inferences.

VISUALIZATION OUTPUT: -> The below bar plot shows us the chart on the findings before.

```
counts <- table(gss$abpoor, gss$income06)
barplot(counts, main = "Distribution by income groups and abpoor", xlab =
"Income groups",
    col = c("darkblue", "red"), legend = rownames(counts))
```



Inference:

Hypothesis Statement: Does there appear to be relationship between “abpoor” and total family income?

H0 => “abpoor” and “income06” are independent (do not vary). Ha => “abpoor” and “income06” are dependent (do vary)

Let us check whether the conditions are met: 1) Independence: Sampled observations are independent. -> Yes in this research sample, every sample are independent to each other. So independence is acheived. 2) the samples taken are random. 3) If sampling without replacement we get n which is less than 10% total population in US. 4) each case only contributes to one cell in the table. If we look at the “mytable” output in Part3, we can see that each case contributes to each income groups.

Method to test the Hypothesis: As we have two categorical variables and atleast one of these has level greater than 2, we go for “chi-square independent test”. As already stated, we are going to find the independence on “abpoor” and “income06”. So we go for this method.

INFERENCE: In this test we are going to find the difference in observed and expected counts between various groups and find the chi-square.

The observed counts can be achieved by:

```
mytable <- table(gss$abpoor, gss$income06)
mytable
##
##       Under $1 000 $1 000 To 2 999 $3 000 To 3 999 $4 000 To 4 999
##   Yes           29              26              19              13
##   No            46              38              27              18
##
##       $5 000 To 5 999 $6 000 To 6 999 $7 000 To 7 999 $8 000 To 9 999
##   Yes              15              17              24              48
##   No               26              31              55              58
##
##       $10000 To 12499 $12500 To 14999 $15000 To 17499 $17500 To 19999
##   Yes              75              64              74              48
##   No              125             110              97              75
##
##       $20000 To 22499 $22500 To 24999 $25000 To 29999 $30000 To 34999
##   Yes              58              67             105             123
##   No              114             117             158             138
##
##       $35000 To 39999 $40000 To 49999 $50000 To 59999 $60000 To 74999
##   Yes              95             175             172             213
##   No              167             267             224             288
##
##       $75000 To $89999 $90000 To $109999 $110000 To $129999
##   Yes              182               172                104
##   No               189               156                108
##
##       $130000 To $149999 $150000 Or Over Refused
##   Yes                 82             209     199
##   No                  53             130     238
```

Then we can find the expected count by the formula given in the video lectures.

Now let us find the overall abpoor rate (“yes”) in the sample: 2408/5362 = 44.9%

Then we should find the chi-square, we get 148.67. The degrees of freedom is (no.of rows-1)*(no. of columns - 1) = 25. We cand find the p-value as:

```
pchisq(148.61, 25, lower.tail = FALSE)
## [1] 1.521e-19
```

The p-value seems to be very low, hence we reject the null hypothesis. Hence, we can come to conclusion that “abpoor” and “income06” are dependent to each other.

Conclusion:

From the inferences, we found that the abpoor do vary as the total family income in any family increases. But there are certain aspects which we need to analyse of which there is some dependent relationship. Some columns like education attained, age, etc are the other factors which may force the mother to abort the child rather than the family income. So this has to be considered.

REFERENCES:

Data Citation

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

URL: http://doi.org/10.3886/ICPSR34802.v1

APENDIX:

The data set used in this project is:

```
head(gss, n = 10)
##    caseid year age    sex  race hispanic uscitzn educ paeduc maeduc
speduc
## 1       1 1972  23 Female White     <NA>    <NA>   16     10     NA
NA
## 2       2 1972  70   Male White     <NA>    <NA>   10      8      8
12
## 3       3 1972  48 Female White     <NA>    <NA>   12      8      8
11
## 4       4 1972  27 Female White     <NA>    <NA>   17     16     12
20
## 5       5 1972  61 Female White     <NA>    <NA>   12      8      8
12
## 6       6 1972  26   Male White     <NA>    <NA>   14     18     19
NA
## 7       7 1972  28   Male White     <NA>    <NA>   13     16     12
NA
## 8       8 1972  27   Male White     <NA>    <NA>   16     16     14
NA
## 9       9 1972  21 Female Black     <NA>    <NA>   12     12     12
NA
## 10     10 1972  30 Female Black     <NA>    <NA>   12     10      7
11
##            degree vetyears sei          wrkstat       wrkslf
marital
## 1        Bachelor     <NA>  NA Working Fulltime Someone Else Never
Married
## 2  Lt High School     <NA>  NA          Retired Someone Else
Married
## 3     High School     <NA>  NA Working Parttime Someone Else
Married
## 4        Bachelor     <NA>  NA Working Fulltime Someone Else
Married
## 5     High School     <NA>  NA    Keeping House Someone Else
Married
## 6     High School     <NA>  NA Working Fulltime Someone Else Never
Married
## 7     High School     <NA>  NA Working Fulltime Someone Else
Divorced
## 8        Bachelor     <NA>  NA Working Fulltime Someone Else Never
Married
## 9     High School     <NA>  NA Working Parttime Someone Else Never
Married
## 10    High School     <NA>  NA Working Fulltime Someone Else
Married
##            spwrksta sibs childs agekdbrn           incom16 born parborn
## 1              <NA>    3      0       NA           Average <NA>    <NA>
## 2     Keeping House    4      5       NA     Above Average <NA>    <NA>
## 3  Working Fulltime    5      4       NA           Average <NA>    <NA>
## 4  Working Fulltime    5      0       NA           Average <NA>    <NA>
## 5  Temp Not Working    2      2       NA     Below Average <NA>    <NA>
```

```
## 6              <NA>    1      0       NA           Average <NA>    <NA>
## 7              <NA>    7      2       NA     Above Average <NA>    <NA>
## 8              <NA>    1      0       NA           Average <NA>    <NA>
## 9              <NA>    2      2       NA           Average <NA>    <NA>
## 10 Working Fulltime    7      4       NA Far Below Average <NA>    <NA>
##    granborn income06 coninc          region          partyid polviews
## 1        NA     <NA>  25926 E. Nor. Central     Ind,Near Dem     <NA>
## 2        NA     <NA>  33333 E. Nor. Central Not Str Democrat     <NA>
## 3        NA     <NA>  33333 E. Nor. Central      Independent     <NA>
## 4        NA     <NA>  41667 E. Nor. Central Not Str Democrat     <NA>
## 5        NA     <NA>  69444 E. Nor. Central  Strong Democrat     <NA>
## 6        NA     <NA>  60185 E. Nor. Central     Ind,Near Dem     <NA>
## 7        NA     <NA>  50926 E. Nor. Central     Ind,Near Dem     <NA>
## 8        NA     <NA>  18519 E. Nor. Central     Ind,Near Dem     <NA>
## 9        NA     <NA>   3704 W. Sou. Central  Strong Democrat     <NA>
## 10       NA     <NA>  25926 W. Sou. Central  Strong Democrat     <NA>
##         relig           attend natspac natenvir natheal natcity natcrime
## 1      Jewish      Once A Year    <NA>     <NA>    <NA>    <NA>      <NA>
## 2    Catholic        Every Week    <NA>     <NA>    <NA>    <NA>      <NA>
## 3  Protestant     Once A Month    <NA>     <NA>    <NA>    <NA>      <NA>
## 4       Other             <NA>    <NA>     <NA>    <NA>    <NA>      <NA>
## 5  Protestant             <NA>    <NA>     <NA>    <NA>    <NA>      <NA>
## 6  Protestant      Once A Year    <NA>     <NA>    <NA>    <NA>      <NA>
## 7    Catholic        Every Week    <NA>     <NA>    <NA>    <NA>      <NA>
## 8      Jewish             <NA>    <NA>     <NA>    <NA>    <NA>      <NA>
## 9  Protestant Sevrl Times A Yr    <NA>     <NA>    <NA>    <NA>      <NA>
## 10 Protestant More Thn Once Wk    <NA>     <NA>    <NA>    <NA>      <NA>
##    natdrug nateduc natrace natarms nataid natfare natroad natsoc natmass
## 1     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 2     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 3     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 4     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 5     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 6     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 7     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 8     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 9     <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
## 10    <NA>    <NA>    <NA>    <NA>   <NA>    <NA>    <NA>   <NA>     <NA>
##    natpark confinan conbus conclerg coneduc confed conlabor conpress
## 1     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 2     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 3     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 4     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 5     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 6     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 7     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 8     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 9     <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
## 10    <NA>     <NA>   <NA>     <NA>    <NA>   <NA>     <NA>     <NA>
##    conmedic contv conjudge consci conlegis conarmy joblose jobfind
## 1      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 2      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 3      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 4      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 5      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 6      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 7      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 8      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 9      <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
## 10     <NA>  <NA>     <NA>   <NA>     <NA>    <NA>    <NA>    <NA>
##              satjob richwork jobinc jobsec jobhour jobpromo jobmeans
## 1  A Little Dissat     <NA>   <NA>   <NA>    <NA>     <NA>      <NA>
## 2              <NA>     <NA>   <NA>   <NA>    <NA>     <NA>      <NA>
## 3   Mod. Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>      <NA>
```

```
## 4   Very Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 5              <NA>     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 6   Mod. Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 7   Very Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 8  A Little Dissat     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 9   Mod. Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
## 10  Mod. Satisfied     <NA>   <NA>   <NA>    <NA>     <NA>     <NA>
##             class rank         satfin    finalter       finrela unemp
## 1   Middle Class   NA Not At All Sat      Better       Average  <NA>
## 2   Middle Class   NA   More Or Less Stayed Same Above Average  <NA>
## 3  Working Class   NA       Satisfied      Better       Average  <NA>
## 4   Middle Class   NA Not At All Sat Stayed Same       Average  <NA>
## 5  Working Class   NA       Satisfied      Better Above Average  <NA>
## 6   Middle Class   NA   More Or Less      Better Above Average  <NA>
## 7   Middle Class   NA   More Or Less      Better Above Average  <NA>
## 8  Working Class   NA Not At All Sat      Better       Average  <NA>
## 9  Working Class   NA   More Or Less       Worse       Average  <NA>
## 10 Working Class   NA Not At All Sat Stayed Same Below Average  <NA>
##    govaid getaid union getahead parsol kidssol abdefect abnomore abhlth
## 1    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes       Yes    Yes
## 2    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes        No    Yes
## 3    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes       Yes    Yes
## 4    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>        No        No    Yes
## 5    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes       Yes    Yes
## 6    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes       Yes    Yes
## 7    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes        No    Yes
## 8    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes       Yes    Yes
## 9    <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes        No    Yes
## 10   <NA>   <NA>  <NA>     <NA>   <NA>    <NA>       Yes        No    Yes
##    abpoor abrape absingle abany pillok sexeduc divlaw          premarsx
## 1     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA> Not Wrong At All
## 2      No    Yes      Yes  <NA>   <NA>    <NA>   <NA>     Always Wrong
## 3     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA>     Always Wrong
## 4     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA>     Always Wrong
## 5     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA>  Sometimes Wrong
## 6     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA>  Sometimes Wrong
## 7      No    Yes      Yes  <NA>   <NA>    <NA>   <NA> Not Wrong At All
## 8     Yes    Yes      Yes  <NA>   <NA>    <NA>   <NA>  Sometimes Wrong
## 9      No   <NA>       No  <NA>   <NA>    <NA>   <NA> Not Wrong At All
## 10    Yes    Yes       No  <NA>   <NA>    <NA>   <NA>     Always Wrong
##    teensex xmarsex homosex suicide1 suicide2 suicide3 suicide4 fear
owngun
## 1     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 2     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 3     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 4     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 5     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 6     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 7     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 8     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 9     <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
## 10    <NA>    <NA>    <NA>     <NA>     <NA>     <NA>     <NA> <NA>
<NA>
##    pistol shotgun rifle             news tvhours racdif1 racdif2
racdif3
```

```
## 1    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 2    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 3    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 4    <NA>    <NA>  <NA>        Once A Week      NA    <NA>    <NA>
<NA>
## 5    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 6    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 7    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 8    <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
## 9    <NA>    <NA>  <NA> Less Than Once Wk      NA    <NA>    <NA>
<NA>
## 10   <NA>    <NA>  <NA>           Everyday      NA    <NA>    <NA>
<NA>
##    racdif4 helppoor helpnot helpsick helpblk
## 1     <NA>     <NA>    <NA>     <NA>    <NA>
## 2     <NA>     <NA>    <NA>     <NA>    <NA>
## 3     <NA>     <NA>    <NA>     <NA>    <NA>
## 4     <NA>     <NA>    <NA>     <NA>    <NA>
## 5     <NA>     <NA>    <NA>     <NA>    <NA>
## 6     <NA>     <NA>    <NA>     <NA>    <NA>
## 7     <NA>     <NA>    <NA>     <NA>    <NA>
## 8     <NA>     <NA>    <NA>     <NA>    <NA>
## 9     <NA>     <NA>    <NA>     <NA>    <NA>
## 10    <NA>     <NA>    <NA>     <NA>    <NA>
```