

Relationship of financial situation with the opinion of family income

31-Oct-2015

Introduction:

Does the satisfaction on the financial situation (satfin) of a person have any relationship with the opinion of family income (finrela)?

This is an observational study and we wanted to see whether the satisfaction on their financial situation affected the way they think/feel their income is, when compared with american family income in general.

We feel the satisfaction financially of a person, would make a positive impact on their opinion of family income with american public in general.

Data:

All data within each year (1972,1983,1992,2002,2012) will be studied and each year's data will be compared with other years to see whether the year had any impact on these variables.

This data will be used to establish only a correlation between the variables of interest. The relationship between these variables alone is studied and even though there might be a positive or negative relationship, it might not necessarily mean that one variable affects the other, as in causal.

Data frame would be created for each year of data, with only the required columns like year, sex, year, degree, satfin and finrela.

```
gss_1972 = gss[gss$year=="1972",c("caseid","year","sex","year","degree","satfin","finrela")]
gss_1983 = gss[gss$year=="1983",c("caseid","year","sex","year","degree","satfin","finrela")]
gss_1993 = gss[gss$year=="1993",c("caseid","year","sex","year","degree","satfin","finrela")]
gss_2002 = gss[gss$year=="2002",c("caseid","year","sex","year","degree","satfin","finrela")]
gss_2012 = gss[gss$year=="2012",c("caseid","year","sex","year","degree","satfin","finrela")]
```

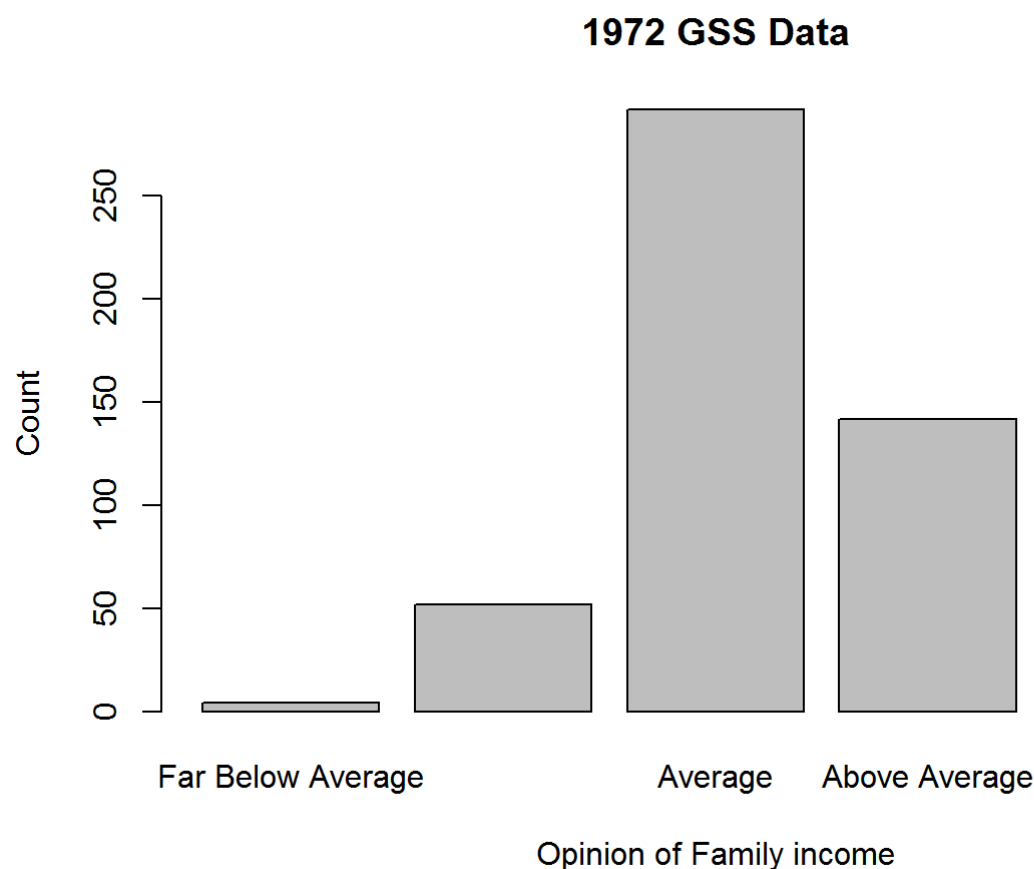
Exploratory data analysis:

This being an observational study, we will use the data subset for each year created from the original data set provided.

Lets first generate the subset of data for those who are satisfied with their financial situation. We believe for people who are satisfied with their financial situation should tend to feel that their opinion of family income is average or above average i.e., in a self content way.

```
gss_1972[complete.cases(gss_1972) & gss_1972$satfin=="Satisfied",] -> satis_1972
gss_1983[complete.cases(gss_1983) & gss_1983$satfin=="Satisfied",] -> satis_1983
gss_1993[complete.cases(gss_1993) & gss_1993$satfin=="Satisfied",] -> satis_1993
gss_2002[complete.cases(gss_2002) & gss_2002$satfin=="Satisfied",] -> satis_2002
gss_2012[complete.cases(gss_2012) & gss_2012$satfin=="Satisfied",] -> satis_2012
```

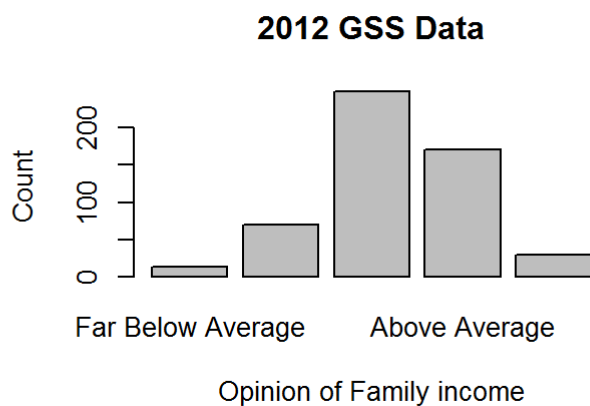
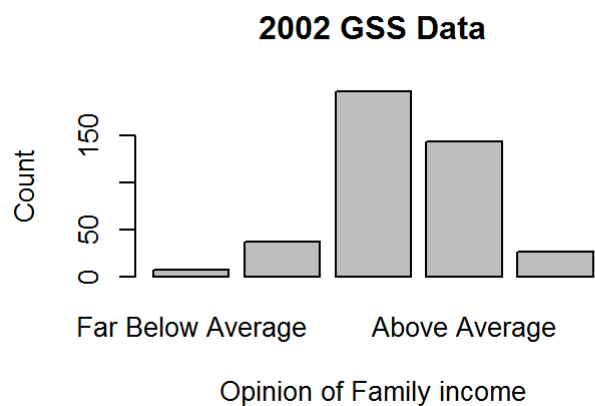
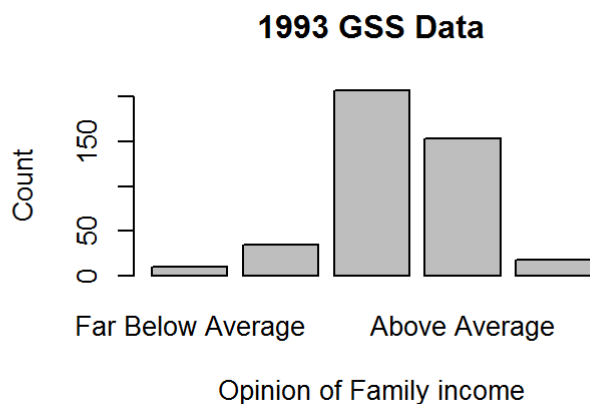
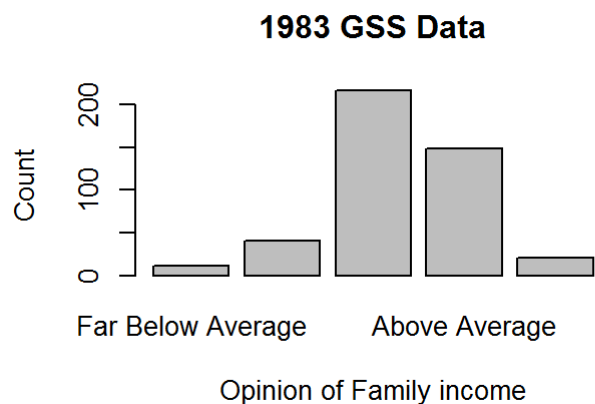
```
plot(satis_1972$finrela,xlab="Opinion of Family income",ylab="Count",main="1972 GSS Data")
```



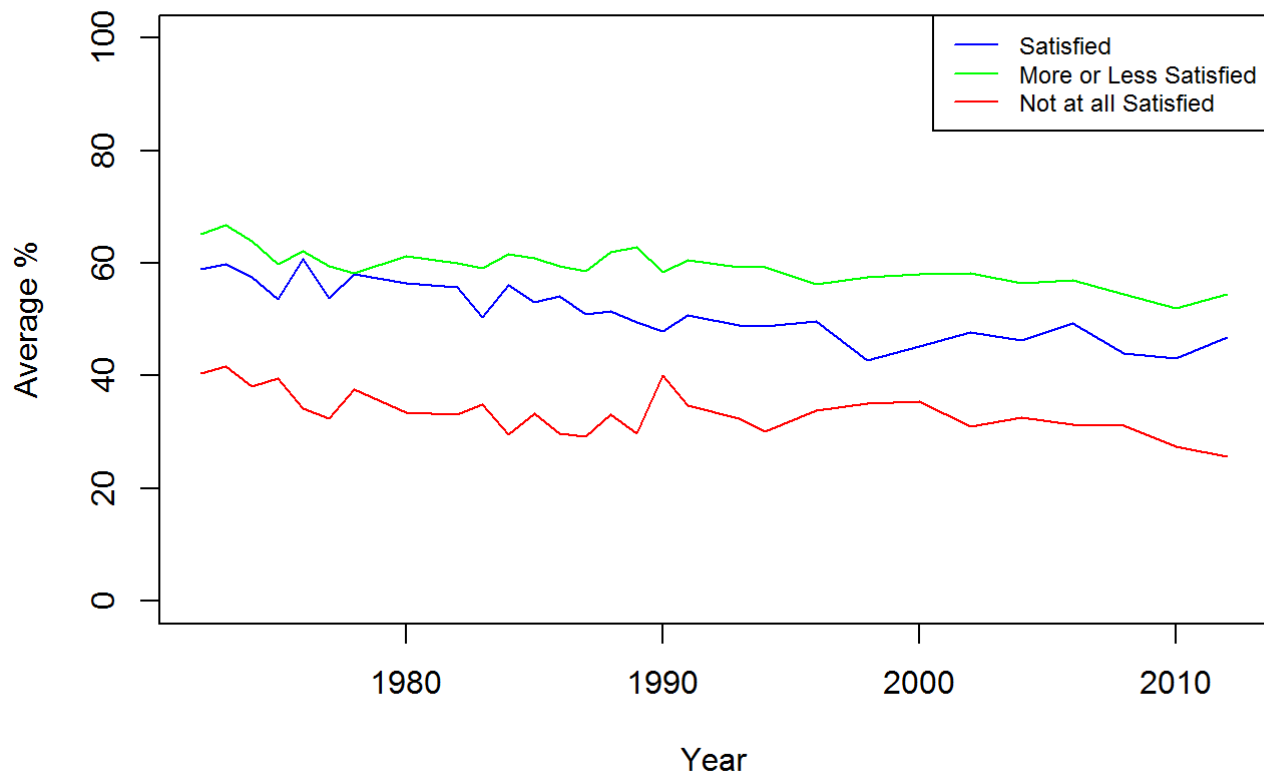
The above plot shows the 1972 GSS data where the people were satisfied with their financial situation. The opinion of family income with general american shows a trend where their opinion being Average and above average, most of the time.

The plot shows a similar trend even for data from the year 1983, 1993, 2002 and 2012.

```
par(mfrow = c(2,2))
plot(satis_1983$finrela,xlab="Opinion of Family income",ylab="Count",main="1983 GSS Data")
plot(satis_1993$finrela,xlab="Opinion of Family income",ylab="Count",main="1993 GSS Data")
plot(satis_2002$finrela,xlab="Opinion of Family income",ylab="Count",main="2002 GSS Data")
plot(satis_2012$finrela,xlab="Opinion of Family income",ylab="Count",main="2012 GSS Data")
```

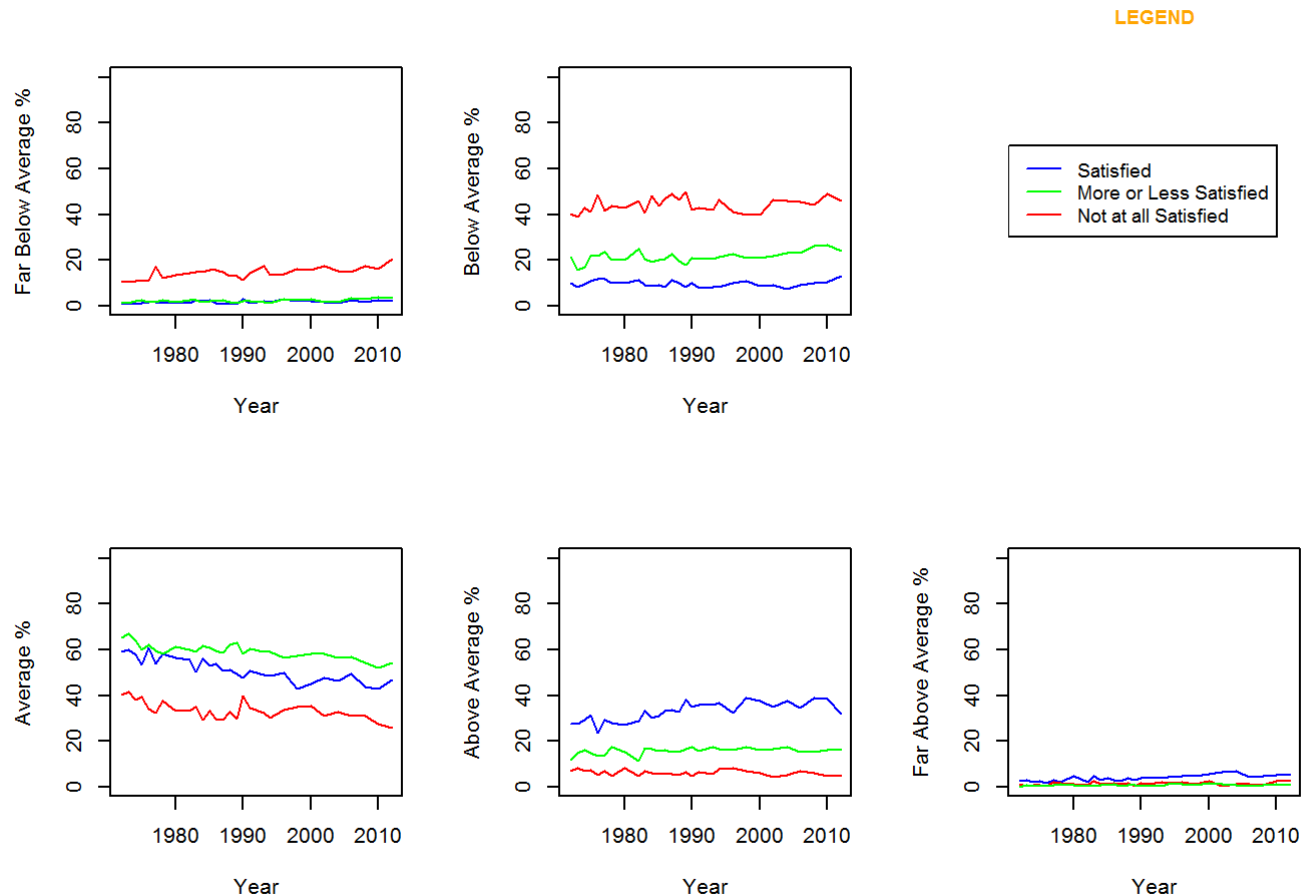


Now let us see the trend of percentage of people over years for each category of Satisfaction on their financial situation with the opinion.



As we see in the above plot, the percentage for “Average” opinion of family income, is very high for those who were either “Satisfied” or “More or Less Satisfied” with their financial situation when compared to those who were “Not at all Satisfied”, for years starting from 1972 until 2012.

Similarly let us create plots for other categories of opinion of family income.



We see that the percentages of people who are “Not at all satisfied” with their financial status have high percentage values for “Below Average” and “Far Below Average”, than the other categories of “Satisfaction of Financial Status”

Similarly people who are “Not at all satisfied” with their financial status have low percentage values for “Average”, “Above Average” and “Far Above Average”, than the other categories of “Satisfaction of Financial Status”, as expected.

We also note that only very few percentage of people have extreme opinions on family income, i.e., for “Far below Average” and “Far Above Average”.

Inference:

Hypothesis: Does the financial satisfaction of a person have any relationship with opinion of family income with american public in general?

Ho (null hypothesis): No. There is no relationship between satfin (satisfaction financially) and finrela (opinion of family income) and they are independent. finrela doesnt vary by satfin. Ha (Alternate hypothesis): satfin (satisfaction financially) and finrela (opinion of family income) are dependent. finrela vary based on satfin.

```
gss_tmp = gss[c("satfin","finrela")]
gss_tmp1 = gss_tmp[complete.cases(gss_tmp),]
summary(gss_tmp1)
```

```
##          satfin          finrela
## Satisfied      :15215  Far Below Average: 2885
## More Or Less   :22980  Below Average   :12575
## Not At All Sat:13816  Average         :25901
##                Above Average    : 9605
##                Far Above Average: 1045
```

```
table(gss_tmp1$satfin,gss_tmp1$finrela)
```

```
##
##          Far Below Average Below Average Average Above Average
## Satisfied                296        1490        7747        5065
## More Or Less              553        4980       13579        3659
## Not At All Sat            2036        6105        4575         881
##
##          Far Above Average
## Satisfied                617
## More Or Less             209
## Not At All Sat           219
```

```
pchisq(11144.3282,8,lower.tail = FALSE)
```

```
## [1] 0
```

Chi-square independence test yielded a very small value (0), thereby rejecting the null hypothesis, which indicates that satfina and finrela are dependent.

Conclusion:

Based on the data analysis done, with the hypothesis testing and with the illustrations (plot) we conclude that the financial satisfaction of a person does have an impact on their opinion of family income with american public in general. We also see from the data and illustrations that it does concur in a positive way.

References:

The following command was used to load the GSS dataset into R, which was used for the analysis
 load(url("http://bit.ly/dasi_gss_data (http://bit.ly/dasi_gss_data)"))

Appendix:

gss_percent - data set

```
head(gss_analysis_dataset,10)
```

##	Year	satfin	finrela_fba%	finrela_ba%	finrela_a%	finrela_aa%
## 1	1972	Satisfied	0.77	10.02	58.96	27.75
## 2	1972	More Or Less	1.69	21.35	65.17	11.80
## 3	1972	Not At All Sat	11.20	40.16	40.44	6.83
## 4	1973	Satisfied	1.31	8.28	59.91	27.45
## 5	1973	More Or Less	1.48	15.68	66.86	15.09
## 6	1973	Not At All Sat	10.70	38.87	41.69	8.17
## 7	1974	Satisfied	0.87	9.78	57.61	29.35
## 8	1974	More Or Less	2.25	17.22	63.92	16.02
## 9	1974	Not At All Sat	10.85	43.11	38.12	7.04
## 10	1975	Satisfied	1.54	10.77	53.63	31.43
##	finrela_faa%		lessthan_avg	greaterthan_avg		
## 1	2.50		10.79	89.21		
## 2	0.00		23.04	76.97		
## 3	1.37		51.36	48.64		
## 4	3.05		9.59	90.41		
## 5	0.89		17.16	82.84		
## 6	0.56		49.57	50.42		
## 7	2.39		10.65	89.35		
## 8	0.60		19.47	80.54		
## 9	0.88		53.96	46.04		
## 10	2.64		12.31	87.70		