

Religion and Religiousness

Introduction:

The research question is: Is there a relation between religion and level of religiousness for the major confessions of the US population?

Religious conflicts and controversy had always been playing an important role in the history of culture. However, in modern culture, we assume, another controversy is becoming more significant: the controversy between religiosity and atheism, between moderate religiousness and fundamentalism. Various occurrences of modern history are showing that the split along the axis of religiousness becomes almost as important as the one along the axis of confessional preferences.

This research can be a part of studies connected with religiousness and confessional controversy.

Data:

1. **Data collection.** Our data are extracted from the General Social Survey database. As the official GSS website proclaims, the vast majority of data is obtained in face-to-face interviews. Under some conditions when it has proved difficult to arrange an in-person interview with a sampled respondent, GSS interviews may be conducted by telephone.
2. **Cases.** Each case represents one respondent who answers the question of GSS.
3. **Variables.** The variables are (according to the official GSS codebook):
 - *relig* - the respondent's religious preference: categorical, nominal (levels: "CATHOLIC", "PROTESTANT" etc.);
 - *relpers* - degree to which the respondent considers himself a religious person: categorical, ordinal (levels: "NOT RELIGIOUS", "SLIGHT RELIGIOUS", "MODRTE RELIGIOUS", "VERY RELIGIOUS").

For the purposes of this study, we suspect religious preferences (*relig*) to be the explanatory variable.

4. **Study.** This is an observational study, since the data were collected in a way that does not directly interfere with how the data arise (i.e. via the survey).
5. **Scope of inference and generalizability.** The population of interest is adults (18+) living in households in the United States. Since this is an observational study based on a random sample, the findings from our analysis can be generalized to the subsection of target population which includes americans belonging to Catholic, Protestant, other Western Christian or Jewish confession.
6. **Causality.** The results of the study cannot be used to establish causal links between the variables of interest since observational studies are based on a random sample (not random assignment) and therefore can establish only association, but not causal links.

Exploratory data analysis:

First, we will inspect our variables one by one.

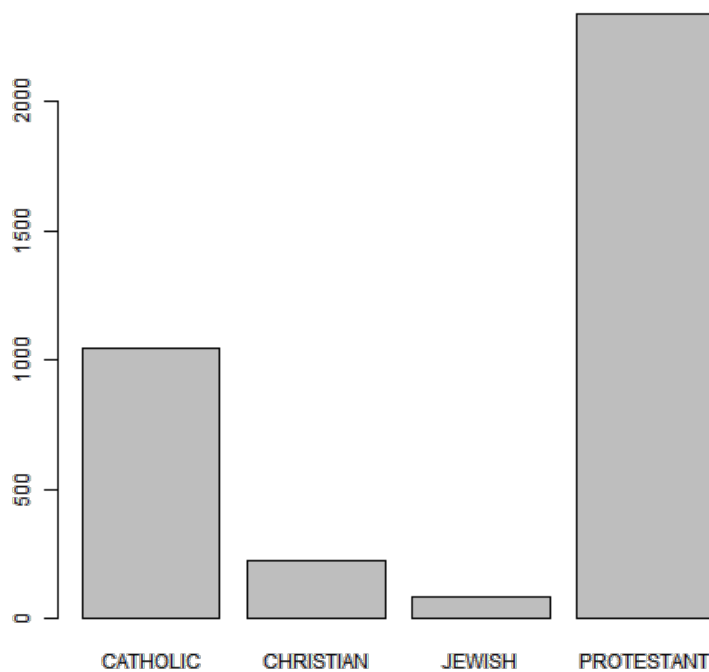
Religious preferences (*relig*):

```
summary(relig)
```

##	CATHOLIC	CHRISTIAN	JEWISH	PROTESTANT
##	1044	225	86	2342

Among the sampled respondents, 2342 are Protestant, 1044 are Catholic, 225 belong to different Western Christian confessions and 86 are Jewish.

```
plot(relig)
```



Level of religiousness (*relpersn*):

```
summary(relpersn)
```

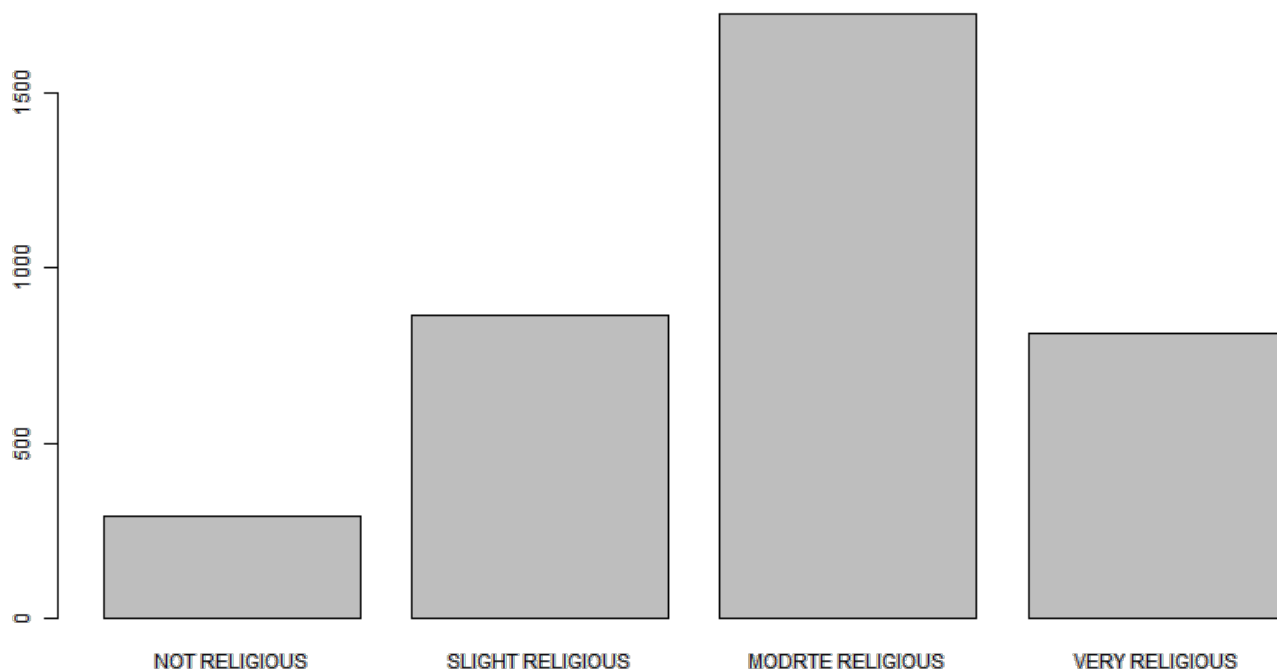
```
##      NOT RELIGIOUS SLIGHT RELIGIOUS MODRTE RELIGIOUS  VERY RELIGIOUS
##              291             864             1727             815
```

```
prop.table(table(relpersn))
```

```
## relpersn
##      NOT RELIGIOUS SLIGHT RELIGIOUS MODRTE RELIGIOUS  VERY RELIGIOUS
##           0.07871           0.23370           0.46714           0.22045
```

Among the total of 3697 respondents, 291 (roughly 7.87%) consider themselves not religious, 864 (23.37%) slightly religious, 1727 (46.71%) moderate religious and 815 (22.05%) very religious. This is our expected proportion for every level of the *relig* variable.

```
plot(relpersn)
```



It looks like the expected proportion of moderate religious respondents is the highest and the expected proportion of not religious is the lowest.

Now let's take a look on how the data are really distributed. We are going to use both contingency table and mosaic plot.

```
table(relig, relpersn)
```

```
##          relpersn
## relig  NOT RELIGIOUS SLIGHT RELIGIOUS MODRTE RELIGIOUS
## CATHOLIC           80           302           529
## CHRISTIAN          28           57           88
## JEWISH             25           30           20
## PROTESTANT        158          475          1090
##          relpersn
## relig  VERY RELIGIOUS
## CATHOLIC           133
## CHRISTIAN           52
## JEWISH              11
## PROTESTANT         619
```

```
prop.table(table(relig, relpersn), margin = 1)
```

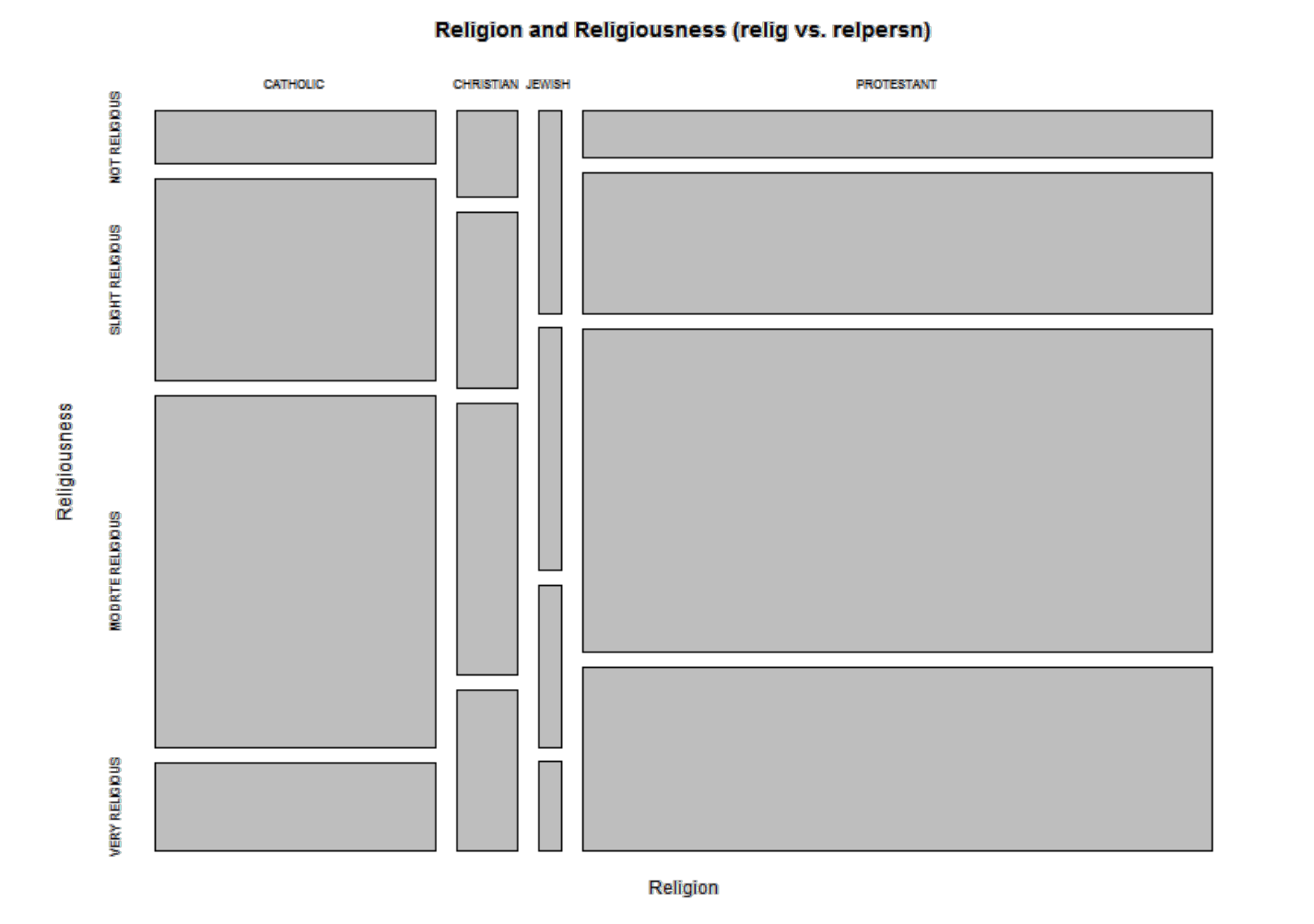
```
##          relpersn
## relig  NOT RELIGIOUS SLIGHT RELIGIOUS MODRTE RELIGIOUS
## CATHOLIC      0.07663      0.28927      0.50670
## CHRISTIAN     0.12444      0.25333      0.39111
## JEWISH        0.29070      0.34884      0.23256
## PROTESTANT    0.06746      0.20282      0.46541
##          relpersn
## relig  VERY RELIGIOUS
## CATHOLIC      0.12739
## CHRISTIAN     0.23111
## JEWISH        0.12791
## PROTESTANT    0.26430
```

Here we can see that the observed data distribution is somewhat different from the expected. For example, the expected number of respondents who are Jewish and who consider themselves not religious is:

```
# the expected proportion of non-religious respondents times the number
# of Jewish respondents
length(relpersn[relpersn == "NOT RELIGIOUS"])/length(relpersn) *
length(relig[relig == "JEWISH"])
```

```
## [1] 6.769
```

Or, in other words, $0.7871247 * 86 =$ roughly 7 cases expected and 25 observed. We won't calculate expected number of observation for each case by hand but rather leave it up for she inference step.



It does look like the proportion is slightly different for different religions. Now we are going to check whether it's true or not using a hypothesis test.

Inference:

1. The hypotheses

H_0 : Religiousness and religious preferences are **independent**. Religiousness rates do not vary by confession.

H_A : Religiousness and religious preferences are **dependent**. Religiousness rates do vary by confession.

2. The method

Since we are going to estimate the relationship between two categorical variables and at least one of them has more than 2 levels, the confidence interval can not be defined. Therefore, we should use **chi-square test of independence**.

Using this method, we should quantify how different the observed counts are from the expected ones. We will find strong evidence against the null hypothesis (i.e. evidence for the alternative hypothesis) if we are able to display large deviations from expected counts. The expected counts are calculated, as we have shown earlier, as a product of observed counts of the dependent variable and expected proportion of an explanatory variable. The difference between observed and expected count is then divided by the expected count. This operation should be performed for every cell of the contingency table and then all derived values are added up together. This is the formula of chi-square.

We may also need to calculate the degrees of freedom (df) for our data distribution. It is calculated as $df = (R - 1) * (C - 1)$, where R stands for the number of rows and C for the number of columns. The df value for our data is:

```
df = (nlevels(relig) - 1) * (nlevels(relpersn) - 1)
df

## [1] 9
```

3. Checking the conditions

I. Independence.

- **random sample/assignment**: we are convinced that GSS uses random sampling; → **met**
- **if sampling without replacement, $n < 10\%$ of population**: 3697, the number of cases, is surely lower than 10% of all adult Americans who are Catholics, Protestant, Christian or Jewish; → **met**
- **each case only contributes to one cell in the table**: only one level for both variables can be chosen during the survey, i.e. the respondent can't at the same time claim himself a Catholic and a Jewish or both moderate religious and very religious person. → **met**

II. Sample size.

Each particular scenario, i.e. cell, must have at least 5 expected cases. In this study we have concentrated our attention only on major confessional groups of the US, thus getting rid of small insignificant values. To prove this condition, we must find the lowest expected count of cases for our data. We do this, accordingly to Bogdanov's Law of minimum, by multiplying the lowest count of the factor (86 cases for the Jewish religion) by the lowest expected proportion of the dependent variable (0.07871 for Not Religious).

```
lowest_count = length(relig[relig == "JEWISH"]) *
(length(relpersn[relpersn ==
"NOT RELIGIOUS"])/length(relpersn))
lowest_count
```

```
## [1] 6.769
```

Our minimum number of expected cases is 6.77 which is greater than 5. → **met**

The conditions are met.

4. The inference

We will perform the calculation for inference using the chi-square function of R.

```
chisq.test(table(relpersn, relig))
```

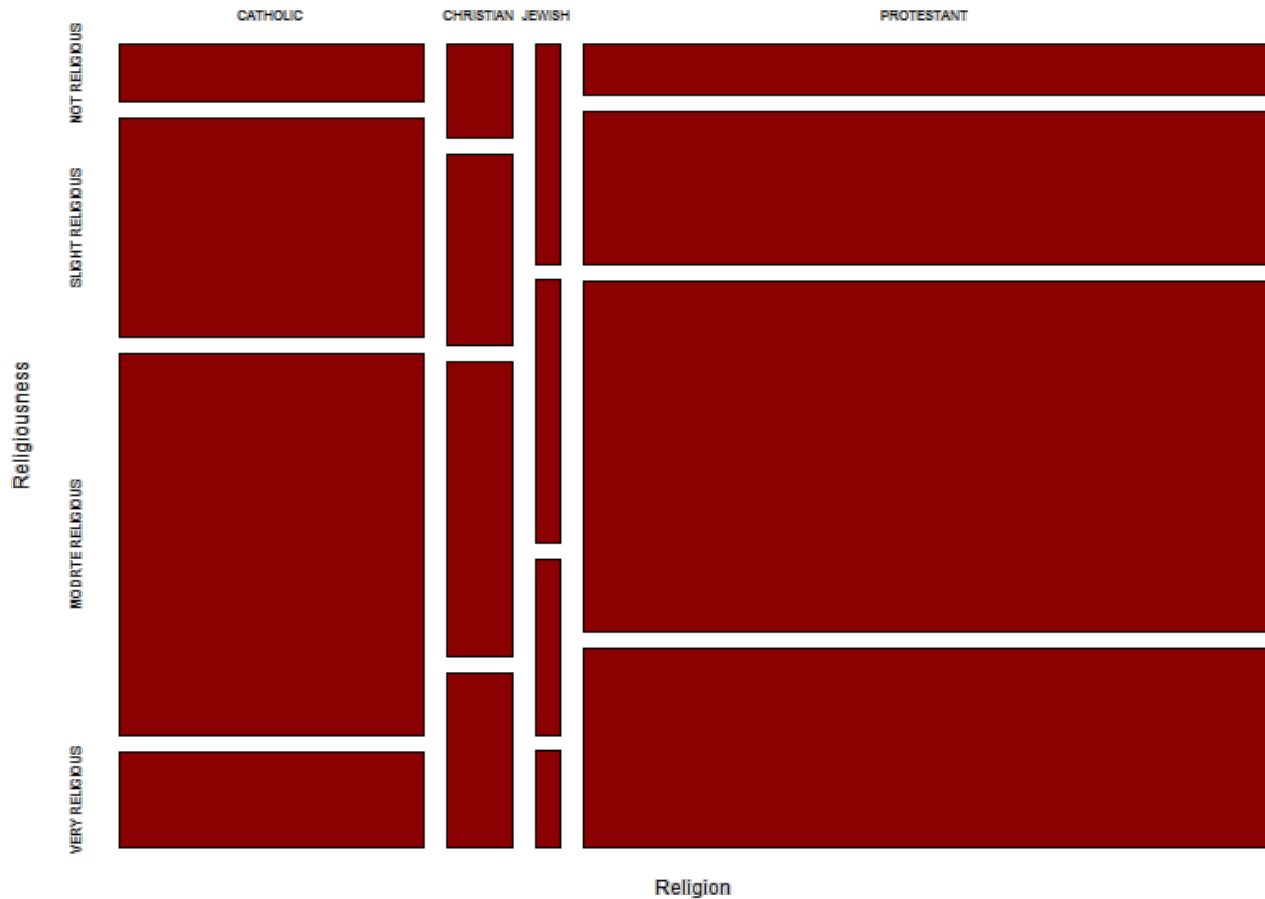
```
##
## Pearson's Chi-squared test
##
## data:  table(relpersn, relig)
## X-squared = 168.8, df = 9, p-value < 2.2e-16
```

Our **p-value** is lower than 0.000000000000000022. With such a tiny p-value we definitely would reject the **H₀** and conclude that the data provide strong evidence for the **alternative hypothesis**.

Therefore, we must assume that ***there is a relation between religion and level of religiousness***. Religion and religiousness are dependent variables and degree of religiousness vary by confession.

Conclusion:

Religion and Religiousness (relig vs. relpersn)



Since we have accepted the alternative hypothesis, we're now able to reflect on what our data is showing. At the first glance, it seems like Jewish people are much more likely to be irreligious than American Christians. It also seems like the Protestants are more likely to be very religious, while the Catholics and the Jews are slightly less. However, this claim should be proved by following studies.

What should have been done better is collecting more cases concerning other confessions so that we could make inference based on that data.

We advisedly excluded respondents who do not associate themselves with any religion (i.e. potential atheists). We did it because the religion variable is assumed to be independent. In case of non-believers the factor and the dependent variable are swapped, so it could disturb the overall picture. My claim is that studies concerning atheism should be designed separately.

The research here provided can be a part of a larger study concerning the religion- or religiosity-based social conflicts.

References:

Citation: Smith, Tom W, Peter Marsden, Michael Hout, and Jibum Kim. *General Social Surveys, 1972-2012* [machine-readable data file] /Principal Investigator, Tom W. Smith; Co-Principal Investigator, Peter V. Marsden; Co-Principal Investigator, Michael Hout; Sponsored by National Science Foundation. –NORC ed.– Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor], 2013.

1 data file (57,061 logical records) + 1 codebook (3,432p.). – (National Data Program for the Social Sciences, No. 21).

Link for download: <http://www3.norc.og/GSS+Website/Download/>

Appendix:

```
datahead = general_ss[1:50, 1:2]
datahead
```

```
##      relig      relpersn
## 1  CATHOLIC SLIGHT RELIGIOUS
## 2  CATHOLIC SLIGHT RELIGIOUS
## 3  PROTESTANT SLIGHT RELIGIOUS
## 4  PROTESTANT NOT RELIGIOUS
## 5  PROTESTANT MODRTE RELIGIOUS
```

##	6	PROTESTANT	VERY	RELIGIOUS
##	7	CHRISTIAN	NOT	RELIGIOUS
##	8	CATHOLIC	MODRTE	RELIGIOUS
##	9	CATHOLIC	MODRTE	RELIGIOUS
##	10	PROTESTANT	SLIGHT	RELIGIOUS
##	11	PROTESTANT	MODRTE	RELIGIOUS
##	12	CATHOLIC	VERY	RELIGIOUS
##	13	PROTESTANT	SLIGHT	RELIGIOUS
##	14	CATHOLIC	MODRTE	RELIGIOUS
##	15	PROTESTANT	VERY	RELIGIOUS
##	16	CATHOLIC	MODRTE	RELIGIOUS
##	17	CATHOLIC	NOT	RELIGIOUS
##	18	CATHOLIC	SLIGHT	RELIGIOUS
##	19	CATHOLIC	VERY	RELIGIOUS
##	20	PROTESTANT	MODRTE	RELIGIOUS
##	21	PROTESTANT	VERY	RELIGIOUS
##	22	PROTESTANT	MODRTE	RELIGIOUS
##	23	CATHOLIC	VERY	RELIGIOUS
##	24	CATHOLIC	MODRTE	RELIGIOUS
##	25	CHRISTIAN	MODRTE	RELIGIOUS
##	26	CATHOLIC	MODRTE	RELIGIOUS
##	27	CATHOLIC	MODRTE	RELIGIOUS
##	28	PROTESTANT	VERY	RELIGIOUS
##	29	PROTESTANT	NOT	RELIGIOUS
##	30	PROTESTANT	MODRTE	RELIGIOUS
##	31	PROTESTANT	MODRTE	RELIGIOUS
##	32	JEWISH	NOT	RELIGIOUS
##	33	CATHOLIC	VERY	RELIGIOUS
##	34	JEWISH	NOT	RELIGIOUS
##	35	JEWISH	MODRTE	RELIGIOUS
##	36	JEWISH	SLIGHT	RELIGIOUS
##	37	JEWISH	NOT	RELIGIOUS
##	38	PROTESTANT	MODRTE	RELIGIOUS
##	39	CATHOLIC	MODRTE	RELIGIOUS
##	40	CATHOLIC	MODRTE	RELIGIOUS
##	41	JEWISH	SLIGHT	RELIGIOUS
##	42	PROTESTANT	MODRTE	RELIGIOUS
##	43	PROTESTANT	SLIGHT	RELIGIOUS
##	44	CHRISTIAN	MODRTE	RELIGIOUS
##	45	PROTESTANT	VERY	RELIGIOUS
##	46	CHRISTIAN	MODRTE	RELIGIOUS
##	47	PROTESTANT	VERY	RELIGIOUS
##	48	CATHOLIC	MODRTE	RELIGIOUS
##	49	PROTESTANT	VERY	RELIGIOUS
##	50	PROTESTANT	MODRTE	RELIGIOUS