IS606 Data Project

Daniel Dittenhafer

October 19, 2015

```
# load data
dataFile <- "C:/Code/R/IS606-ProbStats/FinalProject/data/WVS_subData.RData"
load(dataFile)
colnames(WVS_Subset) <- c("KeyQuestion",
                          "Year",
                          "CountryWave",
                          "Country",
                          "Age",
                          "HighestEducation",
                          "ReligionImportant")
# Subset to just United States
WVS_US <- WVS_Subset[WVS_Subset$Country == 840, ]
WVS_US$CountryName <- "United States"

summary(WVS_US)
```

```
##   KeyQuestion          Year       CountryWave      Country
##  Min.   :-5.000   Min.   :1995   Min.   :8403   Min.   :840
##  1st Qu.:-4.000   1st Qu.:1999   1st Qu.:8404   1st Qu.:840
##  Median : 5.000   Median :2006   Median :8405   Median :840
##  Mean   : 2.236   Mean   :2004   Mean   :8405   Mean   :840
##  3rd Qu.: 8.000   3rd Qu.:2011   3rd Qu.:8406   3rd Qu.:840
##  Max.   :10.000   Max.   :2011   Max.   :8406   Max.   :840
##       Age        HighestEducation ReligionImportant CountryName
##  Min.   :-1.00   Min.   :-3.000   Min.   :-2.0      Length:6223
##  1st Qu.:33.00   1st Qu.: 4.000   1st Qu.: 1.0      Class :character
##  Median :46.00   Median : 6.000   Median : 1.0      Mode  :character
##  Mean   :47.15   Mean   : 5.771   Mean   : 1.8
##  3rd Qu.:61.00   3rd Qu.: 8.000   3rd Qu.: 2.0
##  Max.   :94.00   Max.   : 8.000   Max.   : 4.0
```

Research question

You should phrase your research question in a way that matches up with the scope of inference your dataset allows for.

How have the values of people from the United States changed over time with regard to science and technology and its positive/negative impact on the world? Is there a significant difference in views at different educational levels?

Variables	Description	
005_203	E234	The world is better off, or worse off, because of science and technology
010_023	S020	Year survey
010_028	S024	Country wave
010_004	S003	Country/region
014_003	X003	Age

Variables	Description	
014_030	X025	Highest educational level attained
001_006	A006	Important in life: Religion

Cases

What are the cases, and how many are there?

Each row in the data set is a case and represents the results of an interview conducted by the World Values Survey with an individual age 18 or above. (World Values Survey Association, 2015)

There are 6223 cases in the subset of data to be used in the project.

```
nrow(WVS_US)
```

```
## [1] 6223
```

Data collection

Describe the method of data collection.

According to the World Values Survey website, the data is collected either through "face-to-face interviews or phone interviews for remote areas." (World Values Survey Association, 2015)

Type of study

What type of study is this (observational/experiment)?

The World Values Survey is an observational study, and this data project will be analyzing a subset of the observational study conducted by the World Values Survey Association.

Data Source

If you collected the data, state self-collected. If not, provide a citation/link.

The data was collected and made available by the World Values Survey Association. (World Values Survey Association, 2014).

The link to the data page is: http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp

Response

What is the response variable, and what type is it (numerical/categorical)?

The response variable is the answer provided to the key question, "Is the world is better off, or worse off, because of science and technology?"

The answers are categorical in nature, but ordinal in their degree of support for the better/worse outcome. The distinct answer values and the description of each value are shown in the table below.

Value	Description
1	A lot worse off

Value	Description
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	A lot better off
-5	Missing; Unknown
-4	Not asked in survey
-3	Not applicable
-2	No answer
-1	Don´t know

Explanatory

What is the explanatory variable, and what type is it (numerical/categorical)?

The explanatory variables I've choosen to analyze are the "Highest educational level attained" value and the "Important in life: Religion" variable. Again, these are categorical variables, but ordinal in terms of level of education and degree of importance of religion. The values and descriptions for each variable follow:

Highest educational level attained

Value	Description
1	Inadequately completed elementary education
2	Completed (compulsory) elementary education
3	Incomplete secondary school: technical/vocational type/elementary education, basic vocational qual.
4	Complete secondary school: technical/vocational type/Secondary, intermediate vocational qualification
5	Incomplete secondary: university-preparatory type/Secondary, intermediate general qualification
6	Complete secondary: university-preparatory type/Full secondary, maturity level certificate
7	Some university without degree/Higher education - lower-level tertiary certificate
8	University with degree/Higher education - upper-level tertiary certificate
-5	Missing; Unknown
-4	Not asked in survey
-3	Not applicable; No formal education
-2	No answer
-1	Don´t know

Important in life: Religion

Value	Description
-5	Missing; Unknown
-4	Not asked in survey
-3	Not applicable
-2	No answer
-1	Don´t know
1	Very important
2	Rather important

Value	Description
3	Not very important
4	Not at all important

Relevant summary statistics

Provide summary statistics relevant to your research question. For example, if you're comparing means across groups provide means, SDs, sample sizes of each group. This step requires the use of R, hence a code chunk is provided below. Insert more code chunks as needed.

```
summary(WVS_US)
```

```
##   KeyQuestion          Year       CountryWave      Country
##  Min.   :-5.000   Min.   :1995   Min.   :8403   Min.   :840
##  1st Qu.:-4.000   1st Qu.:1999   1st Qu.:8404   1st Qu.:840
##  Median : 5.000   Median :2006   Median :8405   Median :840
##  Mean   : 2.236   Mean   :2004   Mean   :8405   Mean   :840
##  3rd Qu.: 8.000   3rd Qu.:2011   3rd Qu.:8406   3rd Qu.:840
##  Max.   :10.000   Max.   :2011   Max.   :8406   Max.   :840
##       Age        HighestEducation ReligionImportant CountryName
##  Min.   :-1.00   Min.   :-3.000   Min.   :-2.0      Length:6223
##  1st Qu.:33.00   1st Qu.: 4.000   1st Qu.: 1.0      Class :character
##  Median :46.00   Median : 6.000   Median : 1.0      Mode  :character
##  Mean   :47.15   Mean   : 5.771   Mean   : 1.8
##  3rd Qu.:61.00   3rd Qu.: 8.000   3rd Qu.: 2.0
##  Max.   :94.00   Max.   : 8.000   Max.   : 4.0
```

References

World Values Survey Association. Fieldwork and Sampling. 2015. URL: http://www.worldvaluessurvey.org/WVSContents.jsp.

— WORLD VALUES SURVEY 1981-2014 LONGITUDINAL AGGREGATE v.20150418. Aggregate File Producer: JDSystems. Madrid SPAIN, 2014. URL: http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp.