Precept 4

Overview of Study

In this exercise, we examine cross-national differences in attitudes towards domestic violence and access to information. We explore the hypothesis that there is an association at an a aggregate level between the extent to which individuals in a country have access to knowledge and new information, both through formal schooling and through the mass media, and their likelihood of condemning acts of intimate partner violence.

This exercise is in part based on Pierotti (2013). You can read more about this research in this Washington Post article.

We use data from the Demographic and Health Surveys, which are a set of over 300 nationally, regionally and residentially representative surveys that have been fielded in developing countries around the world, beginning in 1992. The surveys employ a stratified two-stage cluster design. In the first stage enumeration areas (EA) are drawn from Census files. In the second stage within each EA a sample of households is drawn from an updated list of households. In addition, the surveys have identical questionnaires and trainings for interviewers, enabling the data from one country to be directly compared with data collected in other countries. It is important to note that different groups of countries are surveyed every year.

In the study, the author used these data to show that "women with greater access to global cultural scripts through urban living, secondary education, or access to media were more likely to reject intimate partner violence." The data set is in the csv file `dhs_ipv.csv` in the `data` folder and the variables are:

Name	Description
`beat_goesout`	Percentage of women in each country that think a husband is justified to beat his wife if she goes out without telling him.
`beat_burnfood`	Percentage of women in each country that think a husband is justified to beat his wife if she burns his food.
`no_media`	Percentage of women in each country that rarely encounter a newspaper, radio, or television.
`sec_school`	Percentage of women in each country with secondary or higher education.
`year`	Year of the survey
`region`	Region of the world
`country`	Country

Note that there are two indicators of *attitudes towards domestic violence*: `beat_goesout` and `beat_burnfood`. There are also two indicators of *access to information*: `sec_school` and `no_media`.

Part 1: Study Design

Question 1.1

Is this study best understood as a natural experiment, field experiment, laboratory experiment, or descriptive study? Why?

Answer 1.1

This is a descriptive study because the authors observe the real world without there being a random treatment.

Question 1.2

We are going to assess the extent to which exposure to media and education levels affects feelings of whether women feel violence is justified. Come up with two other possible variables that might also help predict the proportion of women who find violence justified.

Answer 1.2

1. Percentage of women who live in cities.
2. Percentage of goverment positions that are held by women.

Question 1.3

Rather than compare two different outcome variables, `beat_goesout` and `beat_burnfood`, how might we instead combine them into a single measure of feelings about violence towards women? How can we generate a measure that allows us to measure when one is high, but the other is not?

Answer 1.3

We could combine them into a single measure by adding them. We could create a measure that indicates when one is high but the other is not by taking the absolute value of the difference. In both cases we would want to make sure the variables are on the same scale, perhaps by Z-scoring them.

Question 1.4

What is the `unit of analysis`? (Use the google.) Does this data allow us to say that a woman's level of education has some predictive relationship with their feelings about violence against women? Why or why not? If not, what sort of relationship does the data allow us to claim?

Answer 1.4

The unit of analysis is the country. We cannot say that a woman's level of education predicts her feelings about violence against women. Rather, her country's education predicts her country's feelings. To attribute this relationship to individuals would be an example of the ecological fallacy.

Part 2: Analyzing the Data

Question 2.1

Read in and summarize the data. How many observations are there? What in the outcomes `beat_burnfood` and `beat_goesout` might give you worry? In what continents are the missing data most prevalent?

Answer 2.1

There are 151 observations. The variables `beat_burnfood` and `beat_goesout` each have a large number of missing values (31 and 27 respectively). The missing data are most prevalent in Sub-Saharan Africa.

```
dhs <- read.csv('data/dhs_ipv.csv', head = TRUE)
nrow(dhs)
```

```
[1] 151
```

```
summary(dhs)
```

```
       X          beat_burnfood    beat_goesout     sec_school
 Min.   :  1.00   Min.   : 0.10   Min.   : 0.30   Min.   : 3.10
 1st Qu.: 40.50   1st Qu.: 4.50   1st Qu.:11.85   1st Qu.:10.18
 Median : 79.00   Median :11.85   Median :28.10   Median :22.40
 Mean   : 80.53   Mean   :15.04   Mean   :28.60   Mean   :24.40
 3rd Qu.:119.50   3rd Qu.:22.25   3rd Qu.:42.08   3rd Qu.:34.90
 Max.   :160.00   Max.   :64.50   Max.   :82.70   Max.   :74.60
                  NA's   :31      NA's   :27      NA's   :3
    no_media          country          year
 Min.   : 0.80   Peru     :  7   Min.   :1999
 1st Qu.:11.25   Uganda   :  6   1st Qu.:2004
 Median :29.15   Egypt    :  5   Median :2007
 Mean   :28.40   Tanzania:  5   Mean   :2007
 3rd Qu.:43.23   Cambodia:  4   3rd Qu.:2011
 Max.   :86.40   Jordan   :  4   Max.   :2014
 NA's   :13      (Other) :120
                           region
 Asia                         :24
 Latin America                :24
 Middle East and Central Asia:19
 Sub-Saharan Africa           :84
```

```
head(dhs)
```

```
  X beat_burnfood beat_goesout sec_school no_media    country year
1 1           4.4         18.6       25.2      1.5    Albania 2008
2 4           4.9         19.9       67.7      8.7    Armenia 2000
3 5           2.1         10.3       67.6      2.2    Armenia 2005
4 6           0.3          3.1       46.0      6.4    Armenia 2010
5 7          12.1         42.5       74.6      7.4 Azerbaijan 2006
6 8            NA           NA       24.0     41.9 Bangladesh 2004
                         region
1 Middle East and Central Asia
2 Middle East and Central Asia
3 Middle East and Central Asia
4 Middle East and Central Asia
5 Middle East and Central Asia
6                         Asia
```

```
table(dhs$region, is.na(dhs$beat_burnfood))
```

```
                              FALSE TRUE
  Asia                           17    7
  Latin America                  20    4
  Middle East and Central Asia   16    3
  Sub-Saharan Africa             67   17
```

```
table(dhs$region, is.na(dhs$beat_goesout))
```

```
                              FALSE TRUE
  Asia                           19    5
  Latin America                  20    4
  Middle East and Central Asia   16    3
  Sub-Saharan Africa             69   15
```

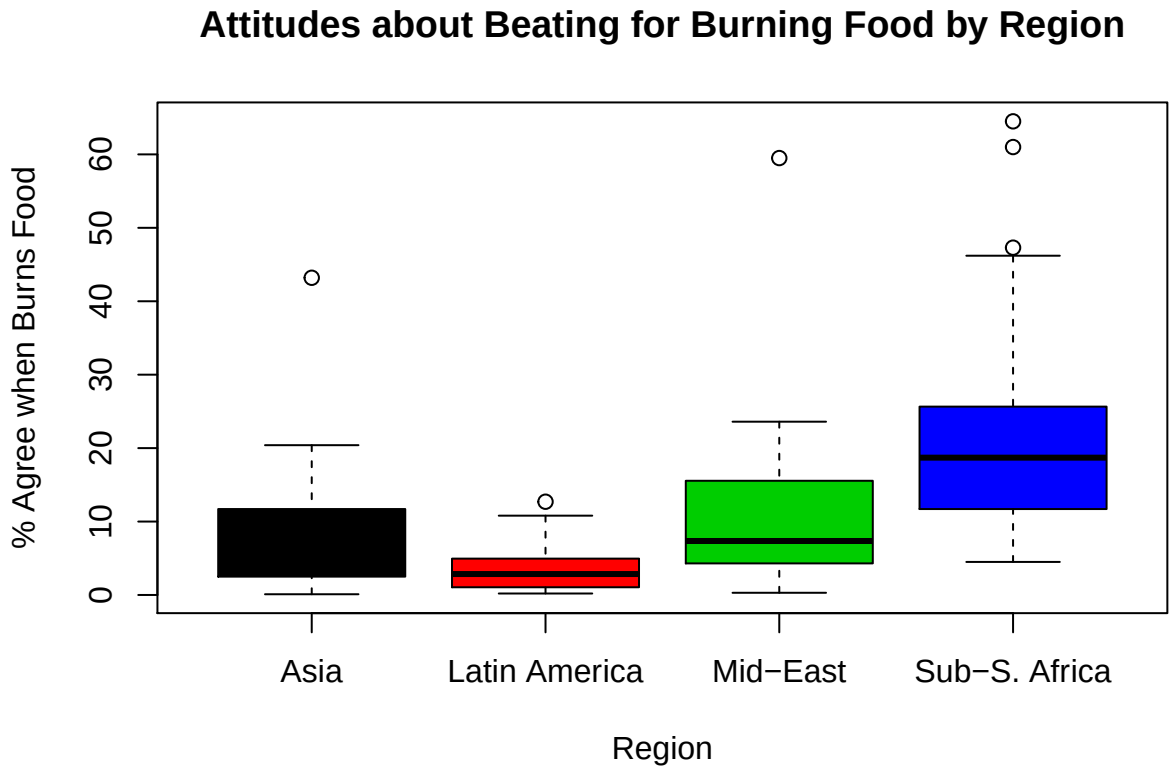
Question 2.2

We first begin by summarizing the outcomes `beat_goesout` and `beat_burnfood` by region. Produce a boxplot for each outcome by region, and summarize the results. Use different colors for each region. Then, create two more boxplots by year. Discuss the results (3-4 sentences, or so).

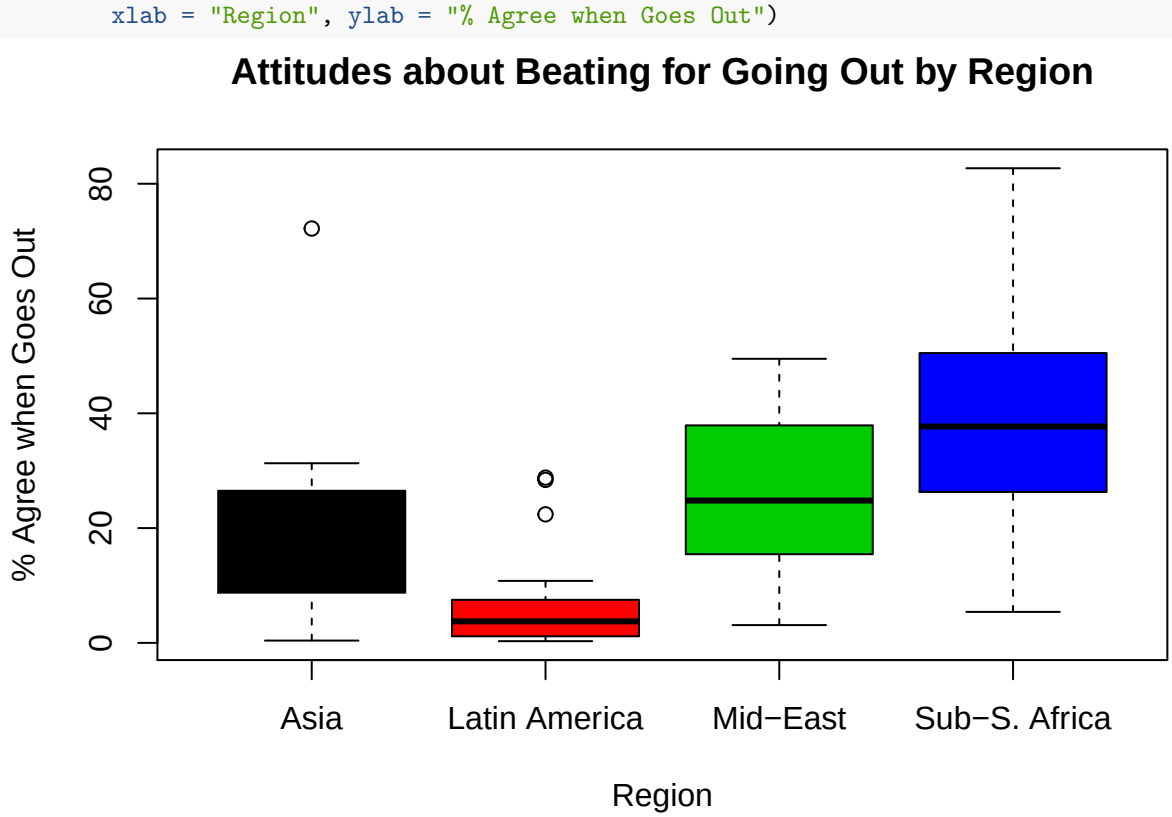
Answer 2.2

Women are generally more supportive of beating women who go out than beating women who burn food. Sub-Saharan African women support both types of beating more than any other group, followed by Middle-East and Central Asian women, then Asian women, then Latin American women. There is not a clear trend over time, although there is some a slight indication that support has decreased in more recent years.

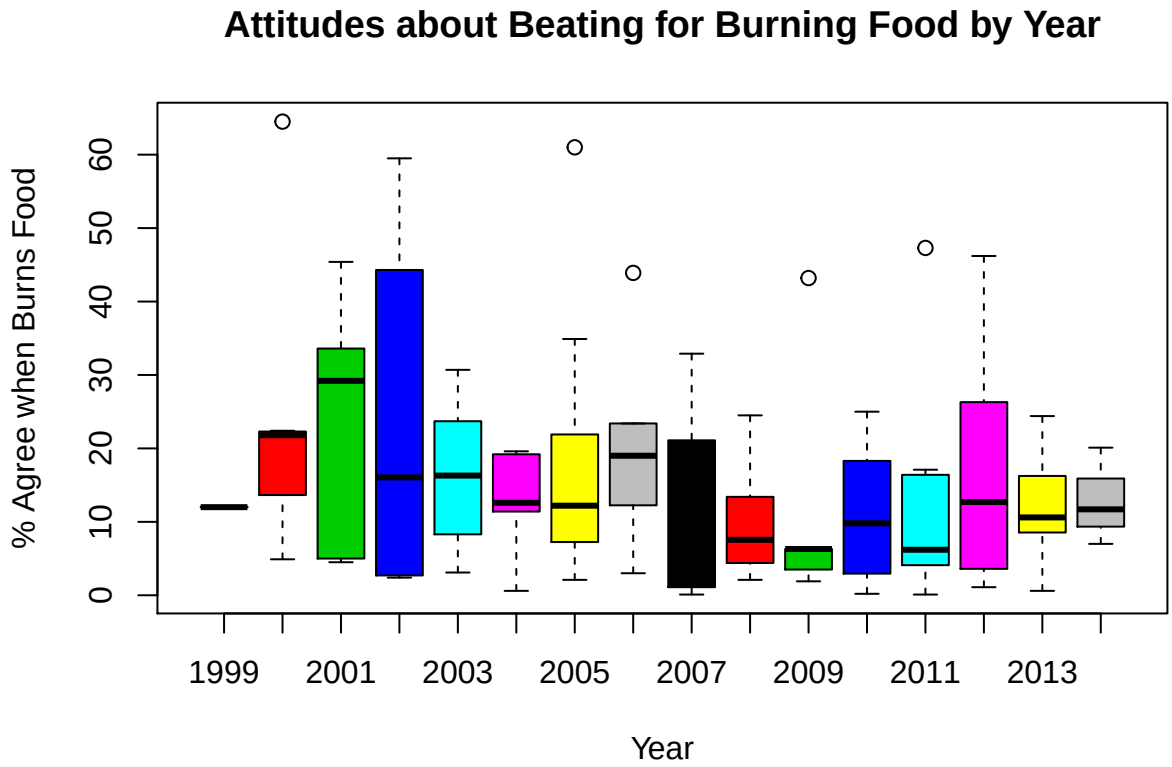
```
# making a new variable with shorter names
dhs$region_short <- c("Asia", "Latin America",
                      "Mid-East", "Sub-S. Africa")[dhs$region]

boxplot(beat_burnfood ~ region_short, data = dhs,
        col = 1:4, xlab = "Region", ylab = "% Agree when Burns Food",
        main = "Attitudes about Beating for Burning Food by Region")
```



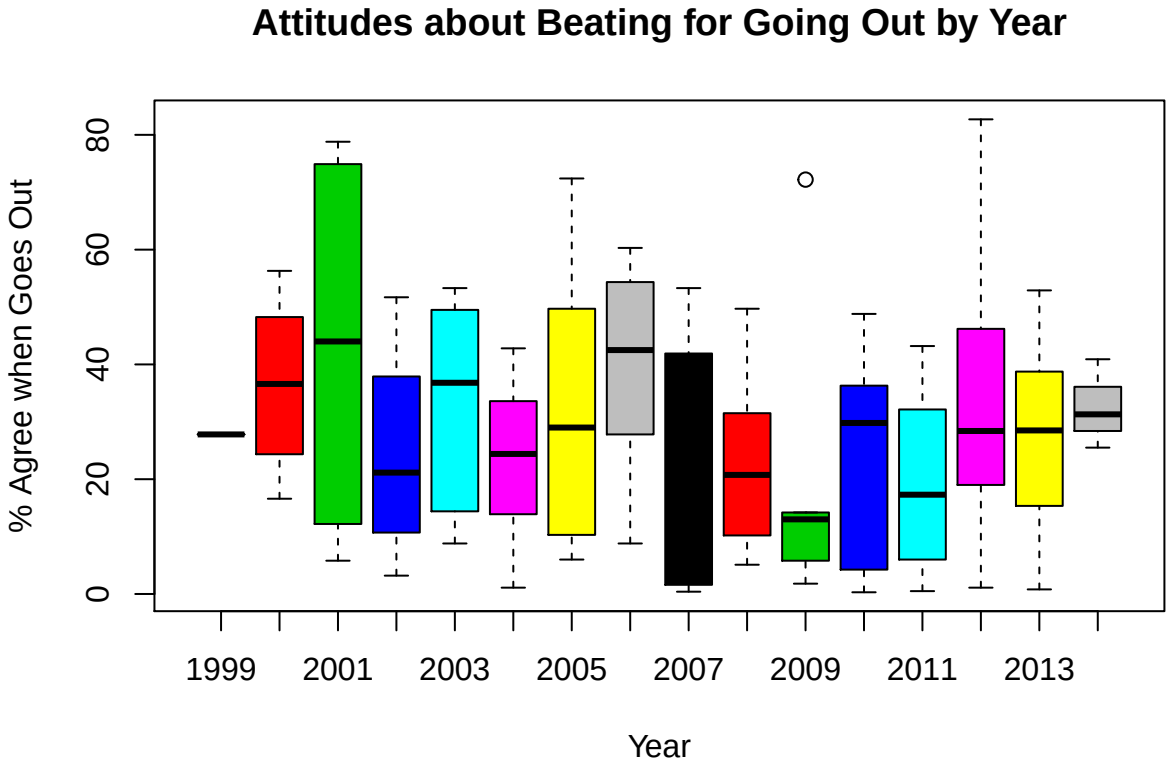

```
boxplot(beat_goesout ~ region_short, data = dhs, col = 1:4,
        main = "Attitudes about Beating for Going Out by Region",
        xlab = "Region", ylab = "% Agree when Goes Out")
```



```
boxplot(beat_burnfood ~ year, data = dhs,
        col = 1:16, xlab = "Year", ylab = "% Agree when Burns Food",
        main = "Attitudes about Beating for Burning Food by Year")
```




```
boxplot(beat_goesout ~ year, data = dhs, col = 1:16,
        main = "Attitudes about Beating for Going Out by Year",
        xlab = "Year", ylab = "% Agree when Goes Out")
```



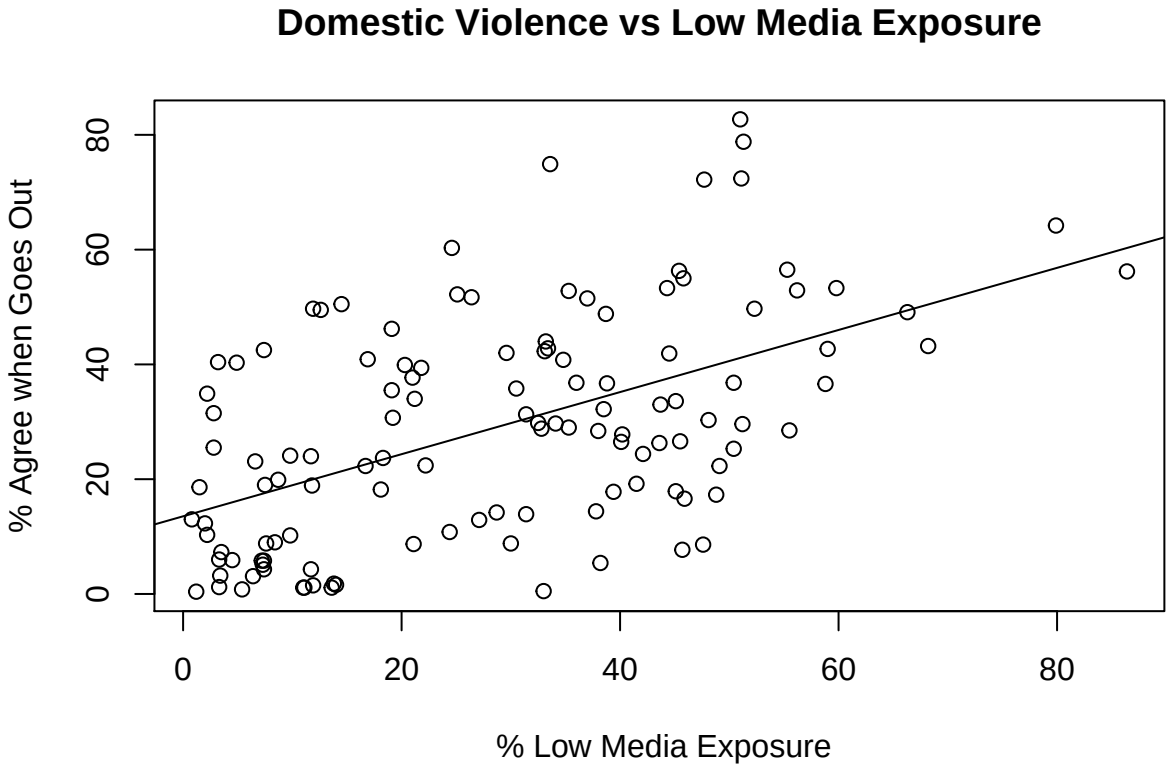
Question 2.3

Now, we turn to analyze the bivariate relationship between `beat_goesout` and `no_media` as well as between `beat_goesout` and `sec_school`. Create a scatterplot for each, and be sure to place the dependent variable on the `y` axis. Add a line of best fit and briefly interpret these graphs in light of the hypothesis of the study.

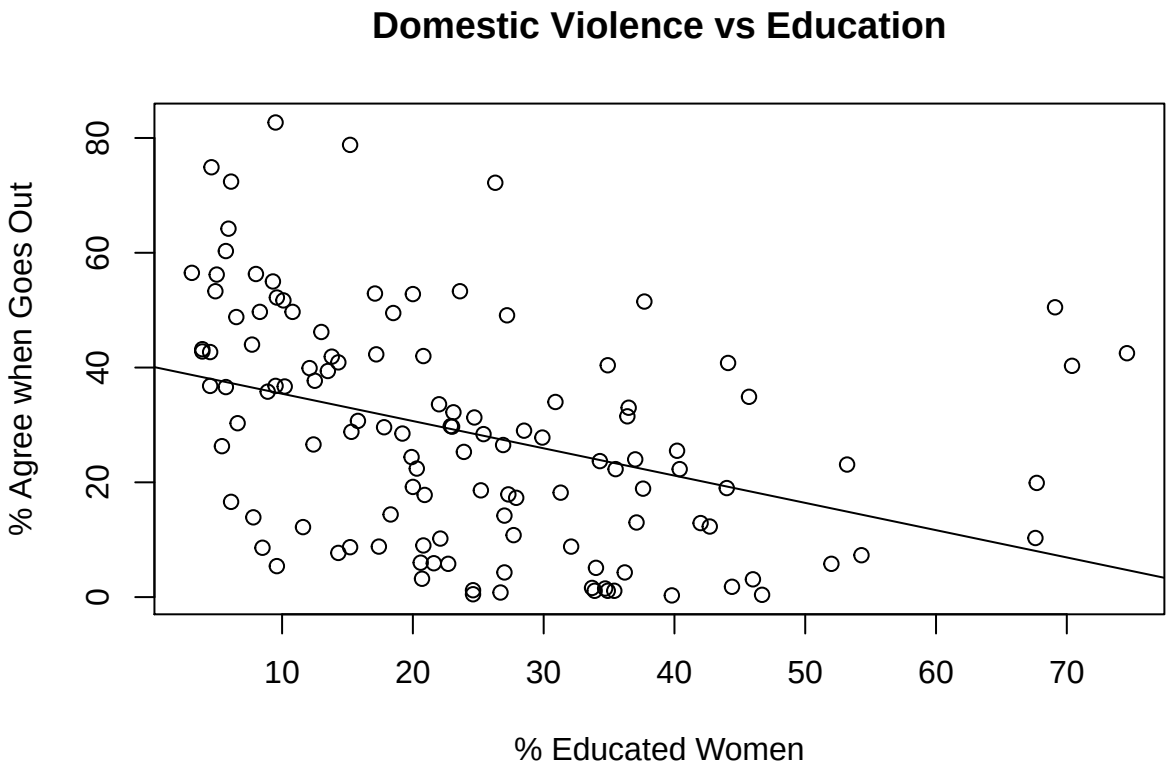
Answer 2.3

Support for domestic violence is associated with low media exposure and low education. This supports the idea that hypothesis that there is a relationship at the country level between access and exposure to new information and support for domestic violence.

```
## Plot 1: Goes Out vs Low Media Exposure
plot(dhs$no_media, dhs$beat_goesout,
     xlab = " % Low Media Exposure", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Low Media Exposure")
abline(lm(beat_goesout ~ no_media, data = dhs))
```




```
## Plot 2: Gooes Out vs Female Education
plot(dhs$sec_school, dhs$beat_goesout,
     xlab = "% Educated Women", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Education")
abline(lm(beat_goesout ~ sec_school, data = dhs))
```



Question 2.4

Report the coefficients from the linear models above, and interpret them. Given the figures above, do the uncovered linear relationships look like they might be systematic, or like they are likely noise?

Answer 2.4

The coefficicents for exposure to media and secondary schooling are .54 and -.47. This means that as the percent of women with low media exposure goes from 80% to 20%, or the percent of women with secondary education goes from 20% to 80%, we would expect to see an associated decrease in support for domestic violence of 32.4 and 28.2 percentage points. This is a large change and the scatterplots show a relatively tight correlation, so I would guess these are systematic covariations and not just noise.

```
lm(beat_goesout ~ no_media, data = dhs)
```

```
Call:
lm(formula = beat_goesout ~ no_media, data = dhs)

Coefficients:
(Intercept)     no_media
    13.5333       0.5412
```

```
lm(beat_goesout ~ sec_school, data = dhs)
```

```
Call:
lm(formula = beat_goesout ~ sec_school, data = dhs)

Coefficients:
(Intercept)   sec_school
    40.1877      -0.4754
```

Question 2.5

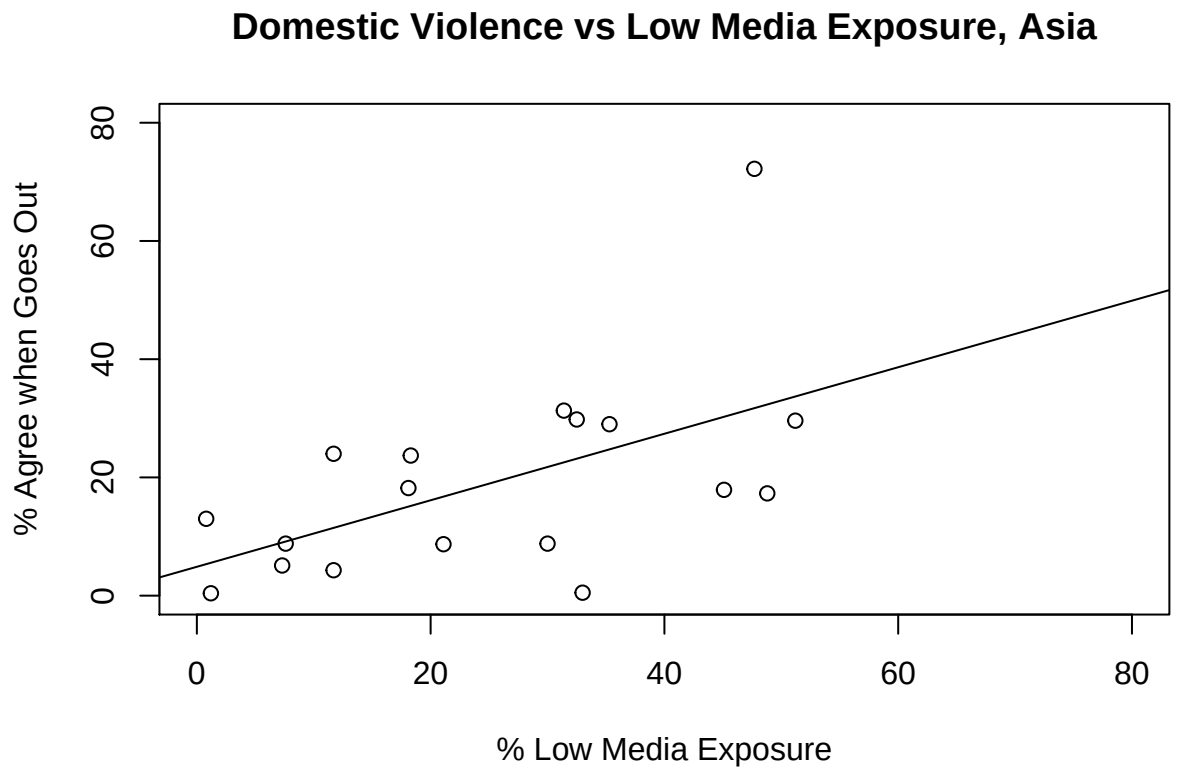
We next revisit the first figure from Question 2.2 using `no_media` and `beat_goesout`. Now, though, create four figures, one for each region. Add the regression lines, as before, and report their values. In which region is the relationship strongest? Weakest?

Answer 2.5

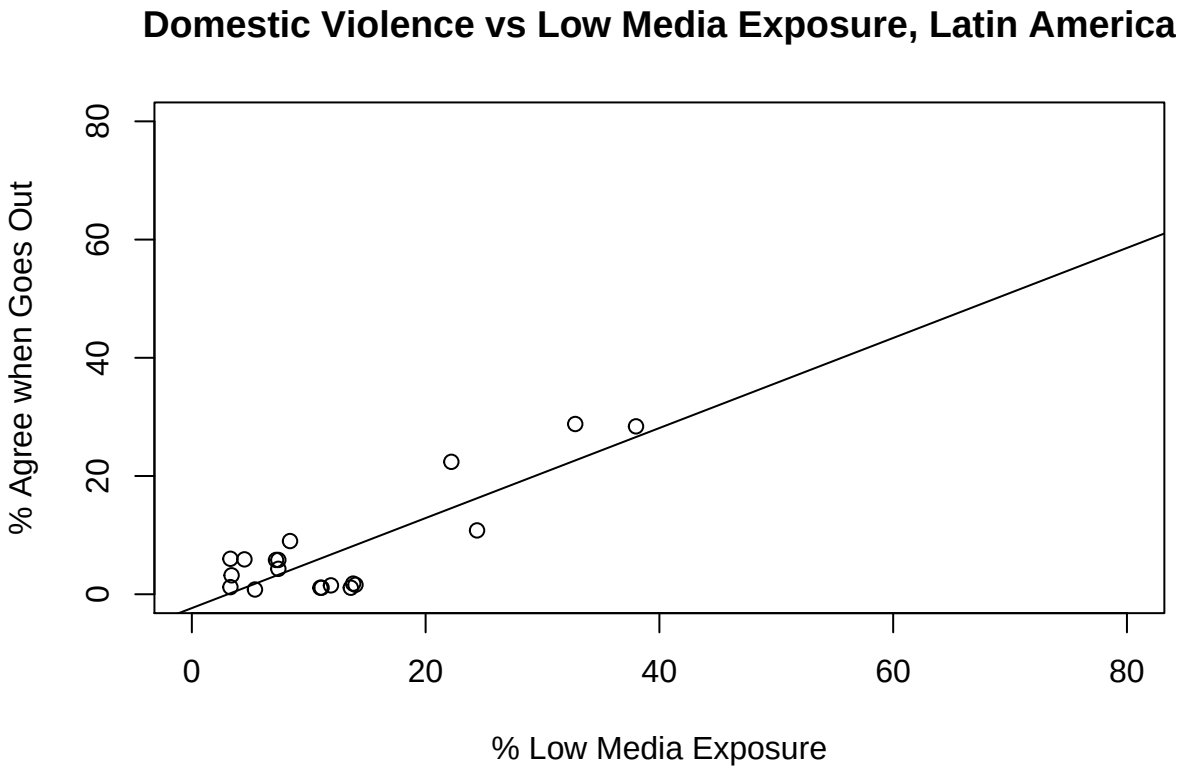
The relationship appears to be strongest in Latin America or the Middle East. While the coefficient is largest for the Middle East, it is also large for Latin America, and the relationship appears to be the tightest in Latin America. The relationship appears to be strongest in Sub-Saharan Africa.

```
m1 <- dhs$region == "Asia"
m2 <- dhs$region == "Latin America"
m3 <- dhs$region == "Middle East and Central Asia"
m4 <- dhs$region == "Sub-Saharan Africa"

plot(dhs$no_media[m1], dhs$beat_goesout[m1],
     xlab = "% Low Media Exposure", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Low Media Exposure, Asia",
```

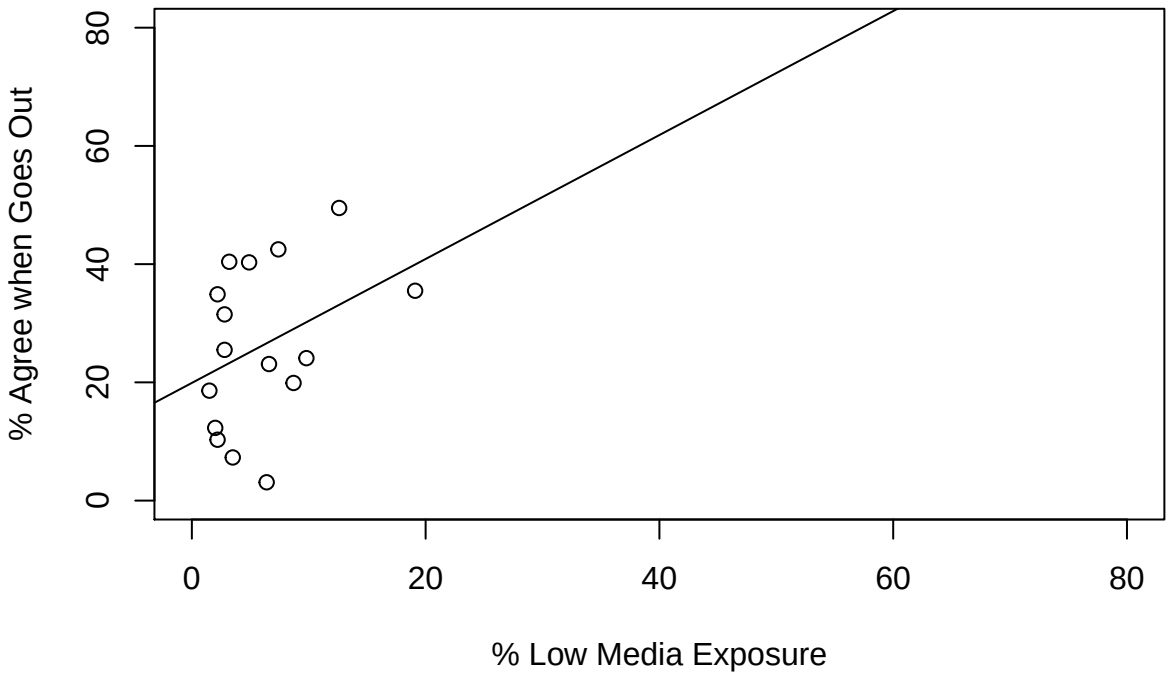
```
     xlim = c(0, 80), ylim = c(0, 80))
abline(lm(beat_goesout ~ no_media, data = dhs, sub = m1))
```



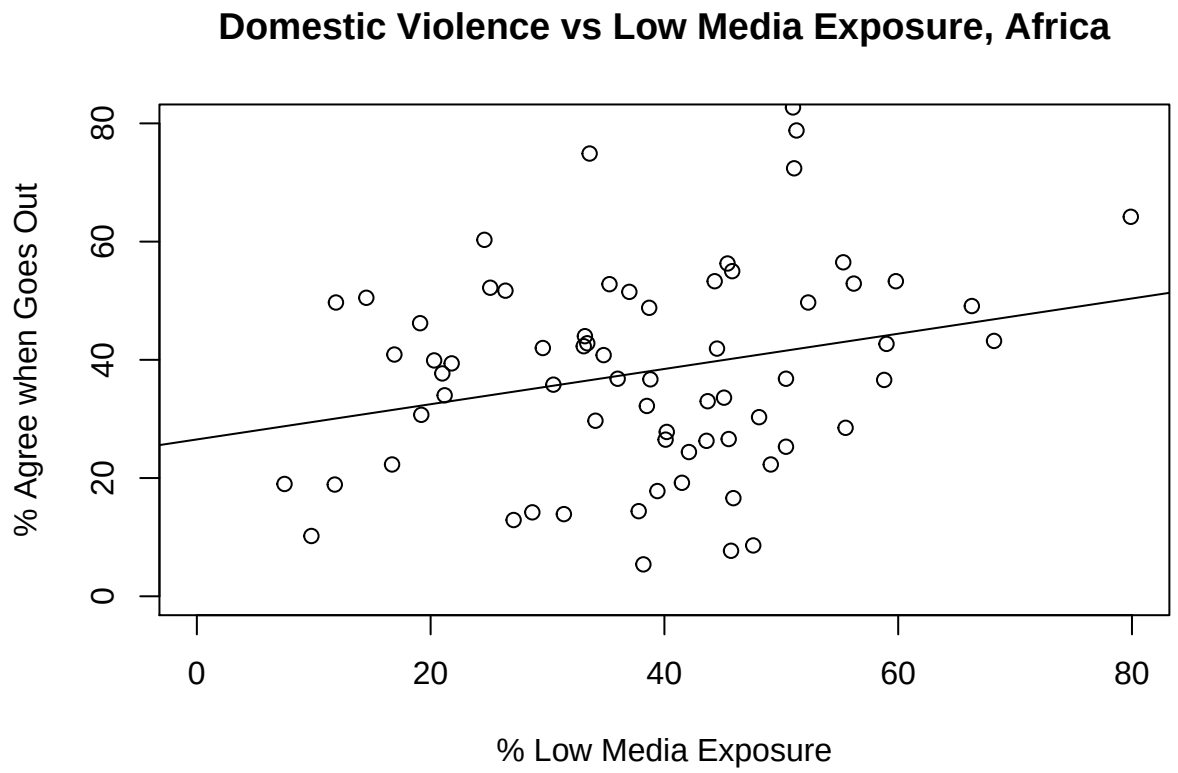

```
plot(dhs$no_media[m2], dhs$beat_goesout[m2],
     xlab = "% Low Media Exposure", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Low Media Exposure, Latin America",
     xlim = c(0, 80), ylim = c(0, 80))
abline(lm(beat_goesout ~ no_media, data = dhs, sub = m2))
```




```
plot(dhs$no_media[m3], dhs$beat_goesout[m3],
     xlab = "% Low Media Exposure", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Low Media Exposure, Mideast",
     xlim = c(0, 80), ylim = c(0, 80))
abline(lm(beat_goesout ~ no_media, data = dhs, sub = m3))
```



```r
plot(dhs$no_media[m4], dhs$beat_goesout[m4],
     xlab = "% Low Media Exposure", ylab = "% Agree when Goes Out",
     main = "Domestic Violence vs Low Media Exposure, Africa",
     xlim = c(0, 80), ylim = c(0, 80))
abline(lm(beat_goesout ~ no_media, data = dhs, sub = m4))
```




```r
lm(beat_goesout ~ no_media, data = dhs, sub = m1)
```

```
Call:
lm(formula = beat_goesout ~ no_media, data = dhs, subset = m1)

Coefficients:
(Intercept)     no_media
     4.8789       0.5627
```

```r
lm(beat_goesout ~ no_media, data = dhs, sub = m2)
```

```
Call:
lm(formula = beat_goesout ~ no_media, data = dhs, subset = m2)

Coefficients:
(Intercept)     no_media
    -2.3444       0.7616
```

```r
lm(beat_goesout ~ no_media, data = dhs, sub = m3)
```

```
Call:
lm(formula = beat_goesout ~ no_media, data = dhs, subset = m3)
```

```
Coefficients:
(Intercept)     no_media
     19.904        1.048
```

```
lm(beat_goesout ~ no_media, data = dhs, sub = m4)
```

```

Call:
lm(formula = beat_goesout ~ no_media, data = dhs, subset = m4)

Coefficients:
(Intercept)     no_media
    26.5303       0.2981
```

References

Pierotti, Rachael S. 2013. “Increasing Rejection of Intimate Partner Violence: Evidence of Global Cultural Diffusion.” *American Sociological Review* 78 (2): 240–65. doi:10.1177/0003122413480363.