Problem Set 2, Part II

Alin Ierima

Access to Information and Attitudes towards Intimate Partner Violence

Hint: read the description carefully!

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 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, Rachel. (2013). "Increasing Rejection of Intimate Partner Violence: Evidence of Global Cultural Diffusion." *American Sociological Review*, 78: 240-265.

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`. The names and descriptions of 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`.

Question 1 (25 points)

Let's begin by examining the association between attitudes towards intimate partner violence and the two exposure to information variables in our data.

- Load the `dhs_ipv.csv` data set and save it in a data frame called `violence`. *(2 points)*
- Use scatterplots to examine the bivariate relationship between `beat_goesout` and `no_media` as well as between `beat_goesout` and `sec_school`.Be sure to add informative axis labels. *(7 points)*
- Repeat these bivariate graphs between `beat_burnfood` and `no_media`, as well as `beat_burnfood` and `sec_school`. *(7 points)*
- Combine all four plots into one figure. *Hint: For this purpose you can use the function `par()` where the argument `mfrow` is set to be equal to `c(nrows,ncols)` before running the plot functions. For example, to plot 6 plots which are organized in two rows and three columns you can write `par(mfrow = c(2,3))`. *(4 points)*
- Briefly interpret these graphs in light of the hypothesis of the study. What associations if any do you observe? Speculate why do we see these relationships. *(5 points)*

Answer 1

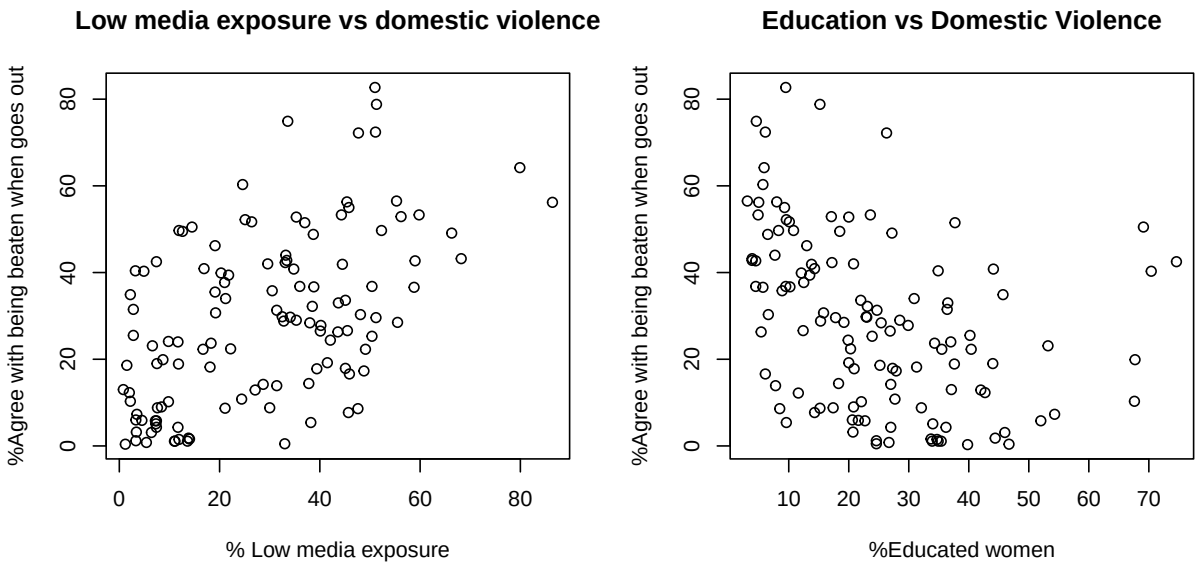
```
library(readr)
violence <- read_csv("data/dhs_ipv.csv")
```

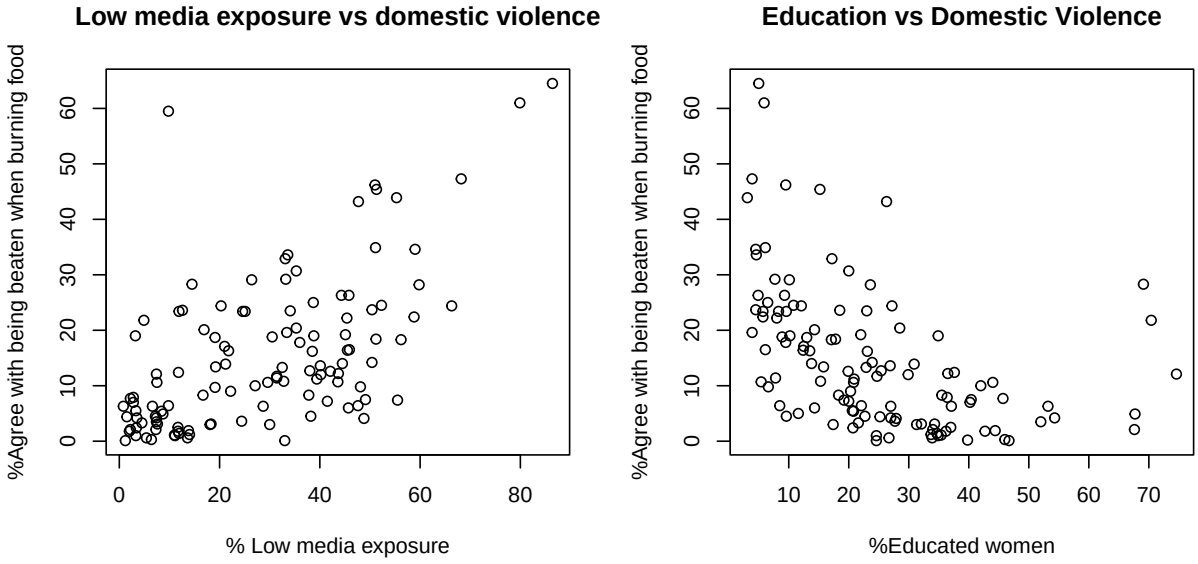
```
## New names:
## Rows: 151 Columns: 8
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (2): country, region dbl (6): ...1, beat_burnfood, beat_goesout, sec_school,
## no_media, year
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...1`
```

Plot: Low Media Exposure vs Domestic Violence

```
## Use this R chunk for your plot. It already specifies the figure size.
par(mfrow = c(2, 2))
plot(violence$no_media,violence$beat_goesout, xlab = "% Low media exposure", ylab="%Agree with being bea
plot(violence$sec_school,violence$beat_goesout, xlab="%Educated women", ylab="%Agree with being beaten


plot(violence$no_media,violence$beat_burnfood, xlab = "% Low media exposure", ylab="%Agree with being b
plot(violence$sec_school,violence$beat_burnfood, xlab="%Educated women", ylab="%Agree with being beaten
```





Insert your answer here: It seems to me that education seems to play a role in the attitude about domestic violence, as the percentage of women supporting it goes down as they receive more education (this being the case for both going out or burning food). In the case of media exposure, it seems that women with higher media exposure are similarly less likely to support domestic violence. In general, there is more support for domestic violence when going out unannounced than when burning food. This supports the idea that education and exposure to mass media leads to people being more aware of human rights or to how people from other regions (such as the Western world) see domestic violence.

Question 2 (35 points)

Make a *"four dimensional"* plot where you show the association between attitudes towards intimate partner violence and media exposure for different regions and level of education.

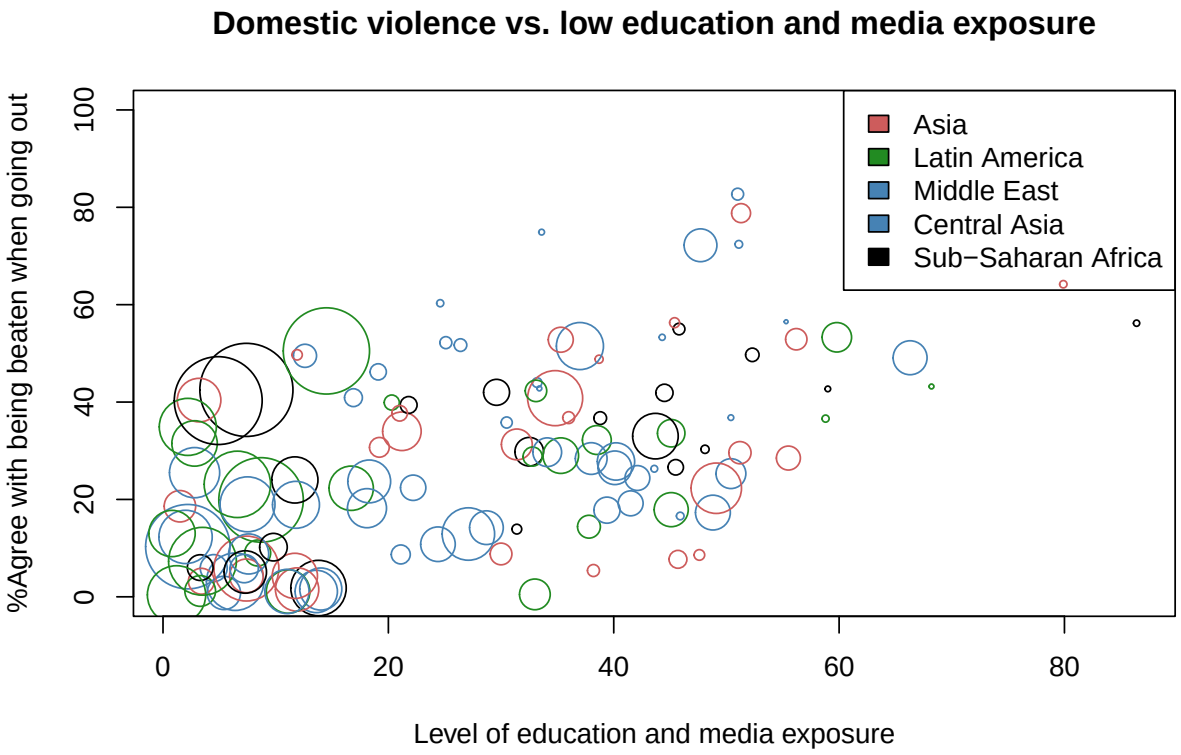
To achieve this do the following steps

- Plot `beat_goesout` against `no_media`, where `no_media` is on the x axis and `beat_goesout` is plotted on the y axis. *(5 points)*
- Color observations from different regions the following way *(5 points)*
  - observations from Asia in `indianred`
  - observations from Latin America in `forestgreen` (*the figure in the answers file is set to a customized green color equal to `"#4C8817"`)
  - observations from Middle East and Central Asia in `steelblue` (*the figure in the answers file is set to a customized blue color equal to `"#2196F3"`)
  - observations from Sub-Saharan Africa in `black`
- Set the size of the circles proportional to years in school education. *(5 points)*
- Make a second figure where you plot `beat_burnsfood` (y axis) against `no_media` (x axis). *(15 points)*
  - Color observations from different regions the same way as above.
  - Set the size of the circles proportional to years in school education - the same way as above.
- Briefly describe in the text the associations that you see in the figures. Why do you think we observe these patterns (these relationships)? *(5 points)*

Answer 2

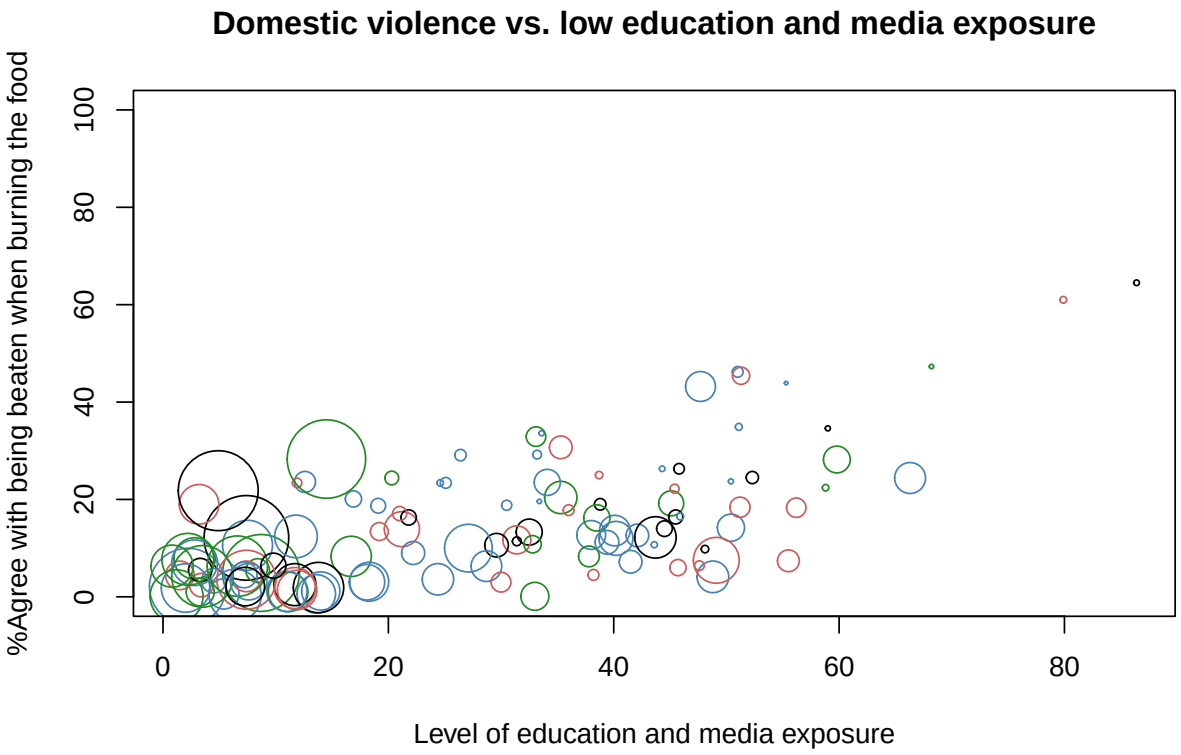
Plot: Goes Out vs Low Media Exposure

```
## Use this R chunk for your plot. It already specifies the figure size.
colors = c("Asia" = "indianred", "Latin America" = "forestgreen", "Middle East" = "steelblue", "Central
colors_region = colors[violence$region]
plot(x = violence$no_media, y= violence$beat_goesout, col=colors, pch=1, cex=violence$sec_school/10, yl
legend("topright", legend = names(colors), fill = colors)
```



Plot: Domestic Violence vs Low Media Exposure and Edication

```
## Use this R chunk for your plot. It already specifies the figure size.
plot(x = violence$no_media, y= violence$beat_burnfood, col=colors, pch=1, cex=violence$sec_school/11, yl
```



Insert your answer here: Largely speaking, it seems that people from the Sub-saharan Africa are the most open to domestic violence and the ones with the least media exposure. On the other side of the spectrum, Latin Americans are the ones with the most media exposure and the lowest acceptance of domestic violence. This further supports the idea of mass media having a role in shaping people's understanding of how acceptable domestic violence is.

Question 3 (10 points)

- Compute the correlation coefficient between `beat_burnfood` and media exposure, as well as between `beat_burnfood` and education. Remember to use complete observations (set the argument `use = "complete.obs"`. *(6 points)*
- What do these measures tell us about the association between education and media exposure with attitudes towards intimate partner violence? *(4 points)*

Answer 3

Correlation between `beat_burnfood` and no media exposure

```
cor(violence$beat_burnfood, violence$no_media, use="complete.obs")
```

```
## [1] 0.5967618
```

Correlation between `beat_burnfood` and education

```r
cor(violence$sec_school, violence$beat_burnfood, use="complete.obs")
```

```
## [1] -0.4760835
```

Insert your answer here: If I understand correctly, the less exposure one has to media, the more they are likely to agree to being beaten for burning food; the effect is similar to the one that more education has. Overall, there are moderately linear relationships between domestic violence and no_media / no education.

Question 4 (30 points)

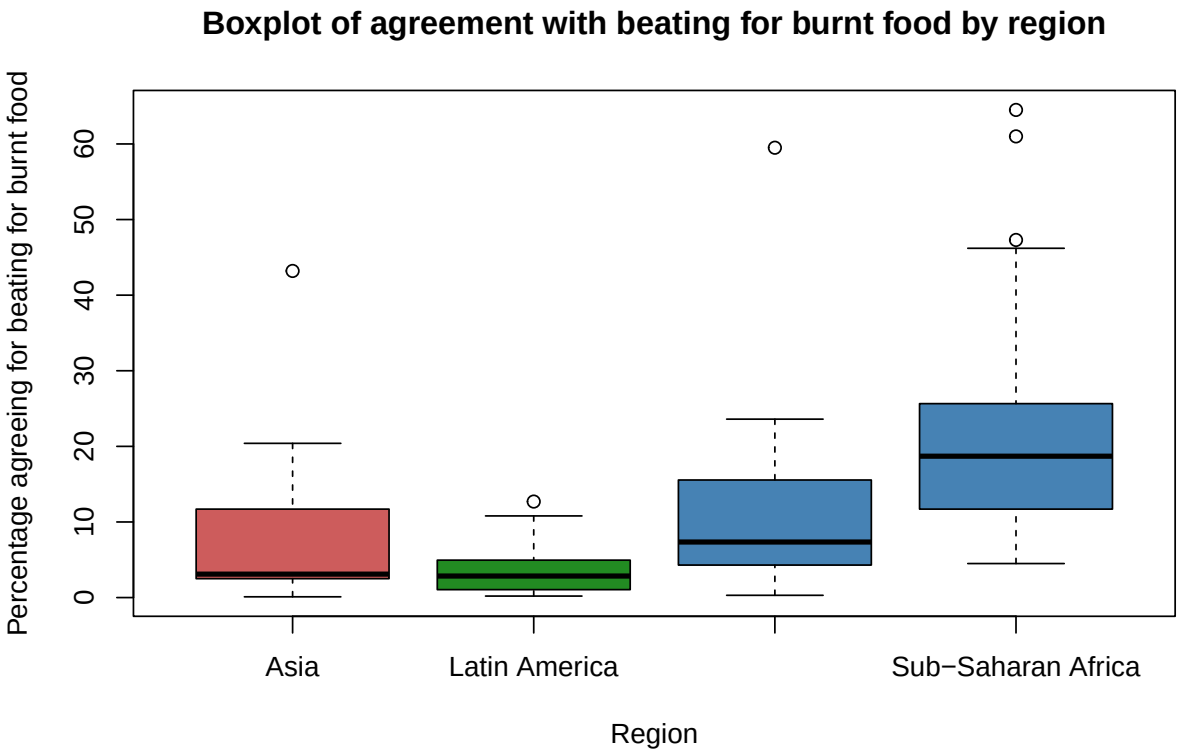
We proceed to explore the national-level differences in attitudes towards domestic violence.

- First, use boxplots to compare the variation in the percentege of `beat_burnfood` between different regions of the world using `region`. *(10 points)*
- What are the main differences across regions in terms of the median and dispersion of the distribution? *(5 points)*
- Second, using boxplots examine the distribution of `no_media` and `sec_school` by region of the world. *(10 points)*
- Comment on the main differences of the distribution of these variables across regions. *(5 points)*

Answer 4

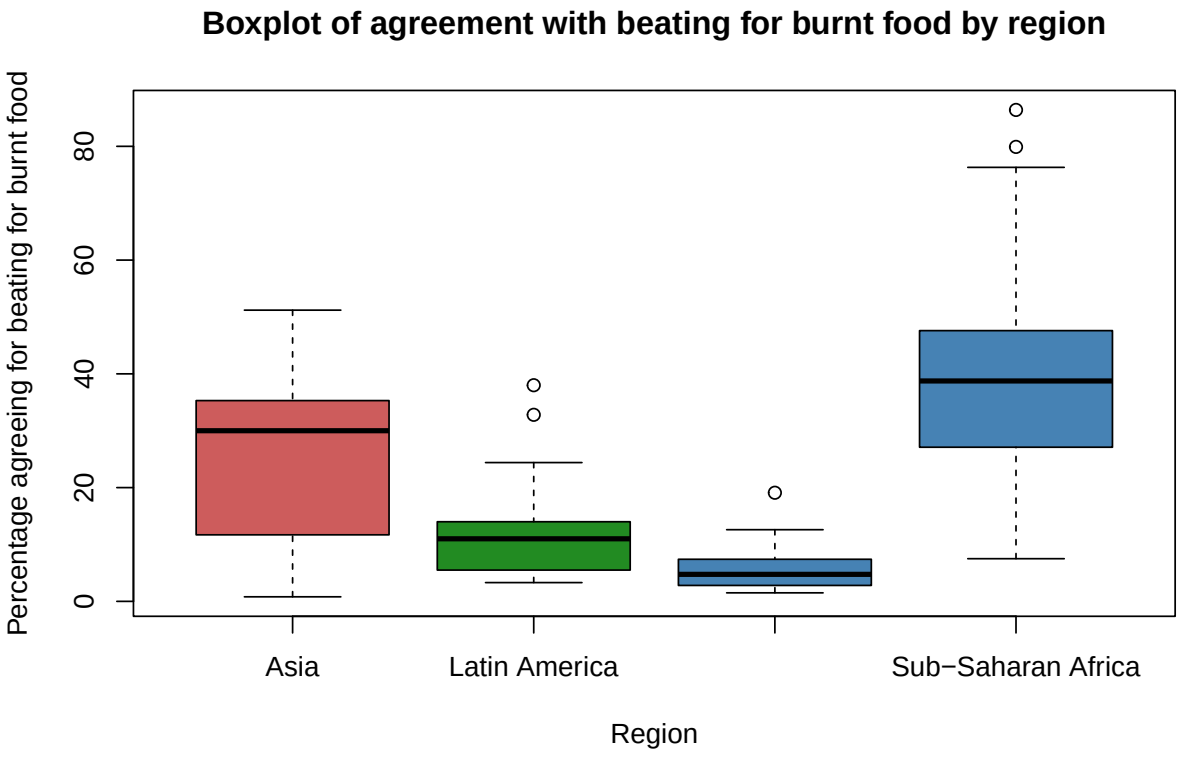
Plot: Domestic Violence Across Regions

```r
## Use this R chunk for your plot. It already specifies the figure size.
boxplot(beat_burnfood~ region, data = violence, col = colors,
        main = "Boxplot of agreement with beating for burnt food by region",
        xlab = "Region", ylab = "Percentage agreeing for beating for burnt food")
```



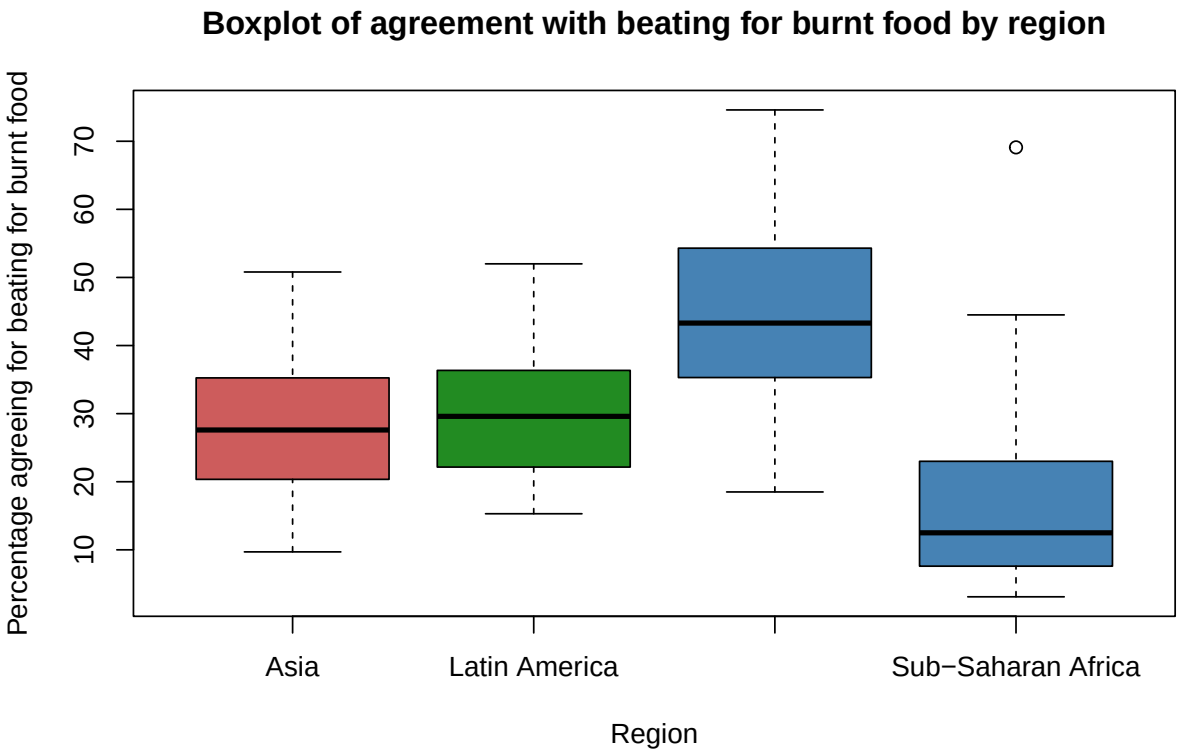
Plot: Low Media Exposure Across Regions

```
## Use this R chunk for your plot. It already specifies the figure size.
boxplot(no_media~ region, data = violence, col = colors,
        main = "Boxplot of agreement with beating for burnt food by region",
        xlab = "Region", ylab = "Percentage agreeing for beating for burnt food")
```



Plot: Female Education Across Regions

```
## Use this R chunk for your plot. It already specifies the figure size.
boxplot(sec_school~ region, data = violence, col = colors,
        main = "Boxplot of agreement with beating for burnt food by region",
        xlab = "Region", ylab = "Percentage agreeing for beating for burnt food")
```



Insert your answer here: The median further proves that sub-saharan Africans agree the most with being beaten when burning food, being at around 17%. Latin Americans have the lowest median, closely followed by Asia (where, interestingly, 50% of the opinions are spread pro-domestically, whereas the anti-domestic violence ones are quite concentrated very very close to the median. Generally, the Middle East and the S-Saharan Africa have the biggest spreads, whereas for Latin America it is quite narrow.

Regarding female education, the IQR is quite similar for all of the regions in terms of size, however the Middle Eastern women are by far the most educated (even considering the women from below the median against the ones from the other regions). Also, the whiskers show us that there is a wide range of education in the M-E, from a minimum of 20% to a maximum of 70%. The minimum coincides with the higher quartile of the sub-saharan african women. Regarding Latin America and Asia, they are quite similar both in regards to their IQR percentage and to their median.

Bonus Question (20 points)

An important point of the researcher's hypothesis is that the support towards intimate partner violence should *decrease* over time, as more women across regions have access to formal schooling and exposure to mass media. To test this idea, using time-series plots, examine the trends in `beat_burnfood` from 1999-2014 *within each region*. **Thinking about the study design, what should we consider before trusting that this plot shows a change over time in attitudes?**

You can follow the following steps:

- Create subsets of the data by regions (variable `violence$region`). (4 points)
- Calculate the mean value of `beat_burnfood` by year for each region and save these values in separate objects. Hopefully, this reminds you of using `tapply`! (5 points)

- Plot the time trend for a given region using the function `plot` (6 points)
  - Set the argument `type = "l"` to plot the values as a line.
  - You can add the time trend for the remaining regions by using the function `points()`, right after your code for the plot. Here you can set the argument `type = "l"` so that R plots a line and not points.
  - You can add the text labels (region names) above each line by using the function `text()`. Within this function you will need to specify the x, y coordinates which indicate where the text should be place. Then you need to specify the text to be plotted, for this purpose set argument `labels = "the text you want to plot"`. To specify the color of the text, set argument `col = "color name"`.
  - The figure in the answers file uses the following color names: `"blue"`, `"green"`, `"purple"`, `"red"`. You can you use any colors of your choice.

- Shortly discuss your findings. Do you trust the time trend you found in the data? Explain why yes, or why not. *(5 points)*

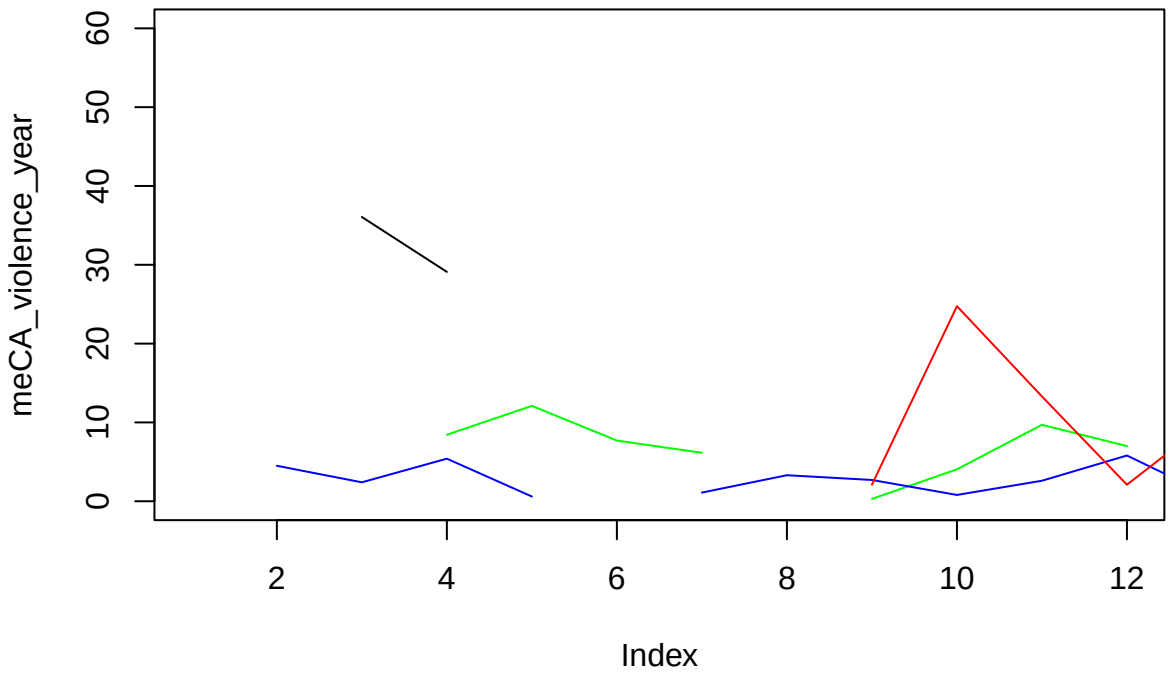
Answer Bonus Question

Plot: Attitudes towards Domestic Violence across Regions, 1999-2014

```
subset_asia = subset(violence, region == "Asia")
subset_meCA = subset(violence, region == "Middle East and Central Asia")
subset_subsahara = subset(violence, region == "Sub-Saharan Africa")
subset_america = subset(violence, region == "Latin America")
 asia_violence_year = tapply(subset_asia$beat_burnfood, subset_asia$year, mean)
meCA_violence_year = tapply(subset_meCA$beat_burnfood, subset_meCA$year, mean)
subsahara_violence_year = tapply(subset_subsahara$beat_burnfood, subset_subsahara$year, mean)
america_violence_year = tapply(subset_america$beat_burnfood, subset_america$year, mean)


plot( meCA_violence_year, type = "l", col = "green", ylim = c(0, 60))

points(subsahara_violence_year, type = "l", col = "black")
points(america_violence_year, type = "l", col = "blue")
points(asia_violence_year, type = "l", col = "red")
```



Insert your answer here:The time trend shows that generally there is a downward trend of the support of domestic violence as time goes by. However, there are big spikes (such as in 2002 for M-E or in 2009 for Asia) that could be caused by the different surveyed countries of that period, and not reflective of the real attitudes for the whole region. Additionally, there are moments of lacking data (such as between 2008-2010 for M-E) which further decrease our trust in the so-called time trend.

Evaluation

- 4 questions for a total of 100 points
- 1 bonus question for 20 points