Case Study 2: Describing and Testing the Association Between Trust in Police and Assault Victimisation in Algeria Using Crosstabulations

Nicholas Tratjenberg-Pareja, Tomáš Diviák

2024-02-05

This activity shows how to undertake bivariate and inferential statistical analysis. The idea is to analyse the relationship between the experience of being victimized and perceptions of trust in public institutions in Africa.

There is a long tradition in criminology of research on public attitudes toward criminal justice and legal institutions given its role on compliance and cooperation with law enforcement (Bradford & Jackson, 2010; Sherman, 2002; Tyler, 2006). One of the key determinants of trust in Criminal Justice Institutions is crime experiences. Previous studies have shown that citizens that were victims of crime are more likely to show more negative perceptions of police and less confidence in them (Berthelot et al., 2018; Callahan & Rosenberg, 2011; O’Connor, 2008; Sprott & Doob, 2009). Alternatively, some research has shown that those that trust less in the police are less likely to report violent crimes (Kääriäinen & Sirén, 2011). Yet, most of the research has been conducted in North America and Europe. Exploring the relationship between victimization and trust in police in the global south is worth considering the elevated levels of violence observed, particularly in Latin America and Africa (UNOCD, 2O23) and also by the under-resourced and weak criminal justice institutions in these regions (Bergman, 2018). Some recent research has shown that trust in police is also higher among citizens that have not been victimized in african countries such as South Africa (Olutola & Bello, 2016), and even that this difference in trust in police due to victimization is a characteristic of developing countries and not so much in more industrialized ones (Alda et al., 2017).

library(here) # to identify the path to the data  
library(tidyverse) # for data transformations etc.  
library(haven) # for importing data  
library(ggplot2) # for visualizations  
library(gmodels) # for crosstables  
library(lsr) # Cramer's V

In this context, we are interested to know if there is a relationship between victimization and trust in a criminal justice institution in Algeria. This exercise will allow us to explore creation of new variables collapsing categories, building cross tables, and analysis of association and testing independence.

In order to explore the relationship between victimisation and trust in criminal justice institutions in Algeria, we will use the Algerian part of the Afrobarometer dataset (Hammani et al., 2015). Using the ‘haven’ package in R (Wickham & Miller, 2020), we will load the dataset and inspect the two variables of interest: having been a victim of assault (named ‘Q11B’ in the original data) and trust in police (‘Q52H’). The ‘table()’ and ‘prop.table()’ functions can be used to inspect the variables of interest.

#Read sav file with survey data  
alg <- read\_sav(here("data/alg\_r6\_data.sav"))  
  
#Inspect variable Q11B, as an example  
table(alg$Q11B)  
prop.table(table(alg$Q11B))

Both the variables have five different possible values, including 0 = not at all, 1 = a little, 2 = somewhat, 3 = a lot, and 9 = NA for confidence in policing’; and 0 = no, 1 = once, 2 = twice, 3 = three or more, and 9 = NA for victimization experiences. In terms of experience with having been assaulted, almost 80% of respondents declared having no such experience in the past year, whereas 14% of respondents claimed to have been victims once with the remaining options not even adding up to 1%. The variable measuring trust in police displays quite an equal distribution of respondents declaring at least some level of trust in police with their share of respondents fluctuating between 25% and 30%, while only 13% percent of respondent declared no trust in police at all. For both variables, values ‘9’ refer to values that we would like to filter out from the analysis. To do this, we use the ‘tidyverse’ suite of R packages (Wickham et al., 2019) and we also simultaneously select only the two variables we will be working with in this example using the pipe operator (‘%>%’).

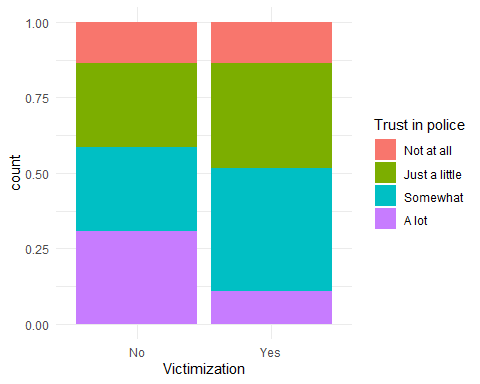
# removing 9s (NAs) from variables  
algS <- alg %>%   
 # select only the columns to be analysed  
 select(Q11B, Q52H) %>%   
 # filter out only valid responses (9 = NA)  
 filter(Q11B != 9, Q52H != 9)  
  
# convert to labelled factors  
algS <- algS %>%   
 mutate(Q11Bf = as\_factor(Q11B),  
 Q52Hf = as\_factor(Q52H))

Since very few respondents suffered more than one assaults, we will create a binary measure of victimization, which distinguishes victims from non-victims.

# collapsing the categories using mutate  
algS <- algS %>%  
 mutate(Q11BfR = recode\_factor(Q11Bf,   
 "No" = "No",   
 "Yes, once" = "Yes",  
 "Yes, twice" = "Yes",  
 "Yes, three or more times" = "Yes"))

A suitable way to visualize the relationship between two categorical variables is to use stacked or clustered barplot. In Figure 3, we use a stacked barplot as implemented in the ggplot2 package (Wickham, 2016). This visualization suggests that the distribution of assault victimization is quite similar across the values of trust in police with the exception of respondents who trust police ‘a lot’, among whom the share of victimised respondents is substantially smaller.

# stacked bar chart  
ggplot(algS, aes(x = Q11BfR, fill = Q52Hf)) +  
 geom\_bar(position = "fill") +  
 labs(x = "Victimization",  
 fill = 'Trust in police') +  
 theme\_minimal()



ggsave(here("exemplar-activities/bar\_graph\_trust.png"),   
 width = 8, height = 6)

Visualization and numeric description of variables can be informative, but it is not sufficient for testing the hypothesis of whether there is an association between these two variables. This is usually done with the use of contingency tables also known as crosstables. The ‘gmodels’ package (Warnes et al., 2019) allows users to create highly informative and visually clear crosstables.

# crosstable with residuals, row percent + a full table  
CrossTable(algS$Q11BfR, algS$Q52Hf, format = c("SPSS"),   
 chisq = TRUE, prop.c = FALSE, prop.chisq = FALSE,   
 prop.t = FALSE, asresid = TRUE)

The crosstabulation (Table 2) displays the categories of trust in police in rows and assault victimization in columns. Each cell contains the number of observations, its row percentage, and its adjusted standardized residual respectively. Below the table, we can also find the information related to Pearson’s chi-squared test of independence. The p-value of this test suggests that we can reject the null hypothesis of independence between these two variables.

We can therefore conclude that the two variables are not independent of each other. Finally, we also used the ‘lsr’ package (Navarro, 2015) to calculate one of the most frequently used coefficients of association for categorical variables – Cramér’s V. Using a coefficient of association allows us to complement the statistical test with a measure of effect size. The resulting value of 0.16 indicates that the relationship between trust in police and being a victim of assault is medium.

# select only the relevant columns  
tab <- table(algS$Q52Hf, algS$Q11BfR)  
# coefficient of association  
tabR <- tab[2:5, 1:2]   
cramersV(tabR)

## [1] 0.155449