

Problem Set 3

Zachary Brandt

2025-11-14

Weather and Witch Killing

1. Construct a new variable for the total number of murders in a village-year (witch + non-witch murders).

```
killings$total_murders <- killings$witch_murders + killings$oth_murders
```

Create a table of summary statistics for all variables in the dataset, including the mean, standard deviation, minimum, maximum, and number of observations, using stargazer, summary or describe commands in R. Discuss any noteworthy patterns. Pay particular attention to the murder and rainfall variables.

```
summary(killings)
```

##	vid	year	witch_murders	oth_murders
##	Min. : 1.00	Min. :1992	Min. :0.00000	Min. :0.00000
##	1st Qu.:17.00	1st Qu.:1994	1st Qu.:0.00000	1st Qu.:0.00000
##	Median :34.00	Median :1997	Median :0.00000	Median :0.00000
##	Mean :35.03	Mean :1997	Mean :0.09103	Mean :0.09103
##	3rd Qu.:54.00	3rd Qu.:2000	3rd Qu.:0.00000	3rd Qu.:0.00000
##	Max. :71.00	Max. :2002	Max. :3.00000	Max. :5.00000
##	any_rain	any_disease	famine	educat
##	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.8571
##	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:3.5263
##	Median :0.0000	Median :0.0000	Median :0.0000	Median :4.2000
##	Mean :0.1712	Mean :0.1481	Mean :0.1739	Mean :4.0346
##	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:4.6875
##	Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :6.6667
##	trad_relig	total_murders		
##	Min. :0.0000	Min. :0.0000		
##	1st Qu.:0.5333	1st Qu.:0.0000		
##	Median :0.6500	Median :0.0000		
##	Mean :0.6541	Mean :0.1821		
##	3rd Qu.:0.8000	3rd Qu.:0.0000		
##	Max. :1.0000	Max. :5.0000		

Now consider the effect of extreme weather on murders in the village.

Install “miceadds” and “sandwich”. Using the `lm.cluster` command, regress total murders (in a village in a particular year) on the indicator for whether a drought or flood occurred in that year. Make sure that error terms should be allowed to be correlated (“clustered”) across years for the same village (use `vid`). Simply use `summary` to report the results in this question. [Note: Results estimated by `lm.cluster` could not be exported directly with `stargazer` so we use `summary` for simplicity. In the section we will teach how to export clustered regression results in a neater way.]

```
library(miceadds)
library(sandwich)
```

```
model <- lm.cluster(data = killing, total_murders ~ any_rain, cluster = "vid")
summary(model)
```

```
## R^2= 0.00125
##
##               Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 0.17377049 0.02208164 7.869457 3.561844e-15
## any_rain     0.04845173 0.04567845 1.060713 2.888203e-01
```

In a second regression, add average years of schooling and proportion of households practicing traditional religions as additional explanatory variables.

```
model <- lm.cluster(data = killing, total_murders ~ any_rain + educat +
                    trad_relig, cluster = "vid")
summary(model)
```

```
## R^2= 0.00738
##
##               Estimate Std. Error    t value    Pr(>|t|)
## (Intercept) 0.3276266665 0.14575336  2.247815515 0.02458796
## any_rain     0.0400209562 0.04263900  0.938599790 0.34793626
## educat      -0.0378694872 0.02644472 -1.432024384 0.15213684
## trad_relig   0.0005751421 0.10422490  0.005518279 0.99559707
```

Finally, consider a possible instrumental variables (IV) approach. Economic theory suggests that extreme economic hardship—such as a famine—may be associated with more violence, including murders. Famine may be caused by extreme rainfall (which would be the instrumental variable).

Write out the first stage regression, the second stage regression, and the reduced form regression.

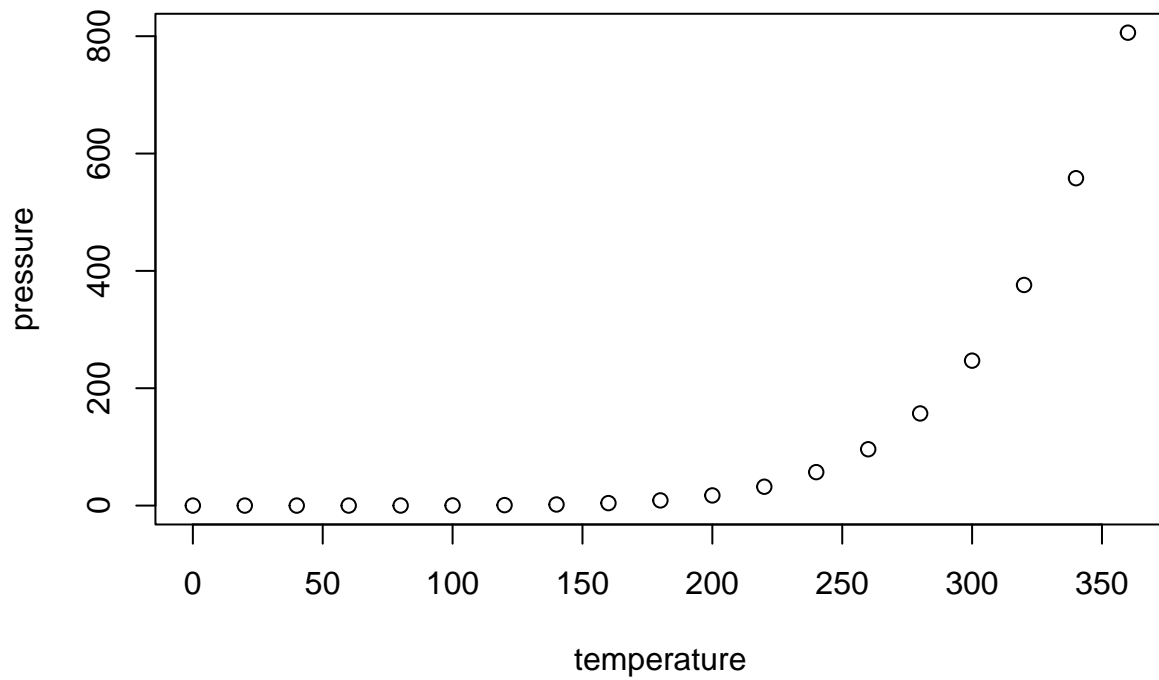
```
# First stage regression
any_rain ~ famine

# Second stage regression
total_murders ~ any_rain + educat + trad_relig

# Reduced form regression
total_murders ~ famine + educat + trad_relig
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.