R.Camp - Homework 2

Santiago López Álvarez

```
library(xtable)
```

```
## Warning: package 'xtable' was built under R version 3.5.2
```

```
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

Homework 2

For this simulation, I am imagining that I could collect all the data of all the conflict-realted deaths in Colombia. In this impossible scenario, I am thinking of having data on the gender of the victims, whether or not they were part of an armed group (regular or irregular), age, and cause of death. If died in combat, I would also know what group the victim was fighting against, and in which area of the country. If this data was available, I could confirm or at least explore with more certainty certain dynamics of the conflict, such as what type of demographic features make joining an armed group more likely to happen.

Variables - Dependent Variable: Number of conflict related deaths in rural areas. Continuous. - Independent Variables: Primary area of residence of the victim (urban or rural). Binary - Discreet. Status of the victim - membership to an armed group or not - (categorical, 4 groups).

- Control Variables: Gender (binary, 1/0), Group that caused the death (categorical, 3 groups).

Simulation

```
rm(list=ls())
set.seed(112) # Set seed for pseudo-random number generator

simN <- 35000    # NUMBER OF TOTAL VICTIMS

# (1-1) GENERATE CORRELATED  VARIABLES
rural <- rbinom(n=simN, size=1, prob=0.7)                  # 70% of probability of living in a rural a
agroup <- rep(NA, simN)
rcivil <- ifelse(rural==1 & runif(simN)<0.65, 1, agroup) # 65% RURAL = CIVILIANS
ccivil <- ifelse(rural==0 & runif(simN)<0.35, 1, rcivil) # 35% OF URBAN = CIVILIANS
armed <- ifelse(is.na(ccivil)==T, 0, 1)                # NON-URBAN = MEMBER OF ARMED GROUP

sim.data <- data.frame(rural, armed)
head(sim.data)
```

```
##   rural armed
## 1     1     0
## 2     0     0
## 3     0     0
## 4     0     1
## 5     1     1
## 6     1     0
```

```
sum(rural)
```

```
## [1] 24516
```

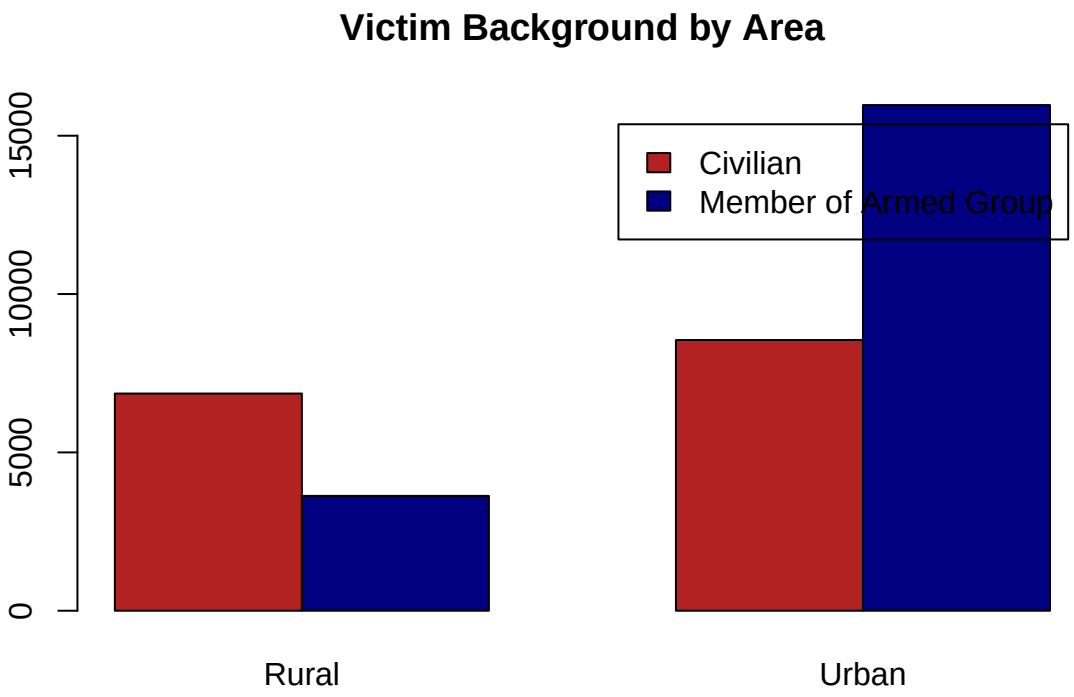
```
sum(armed)
```

```
## [1] 19594
```

```
counts <- t(table(sim.data$rural, sim.data$armed))
counts
```

```
##
##         0     1
##   0  6858  8548
##   1  3626 15968
```

```
par(mfrow=c(1,1))
barplot(counts, main="Victim Background by Area", xlab="", beside=T,
        col=c("firebrick", "navy"), legend=c("Civilian", "Member of Armed Group"),
        axisnames=T, names.arg=c("Rural", "Urban"))
```



```
rho <- cor(rural, armed)
rho   # CORRELATION OF AREA AND MEMBERSHIP TO AN ARMED GROUP
```

```
## [1] 0.2818704
```

```
rm(agroup, rcivil, ccivil, rho, rural, armed)
```

Summary of the data

```
library("stargazer")

mod_stargazer <- function(...){
        output <- capture.output(stargazer(...))
        output <- output[4:length(output)]
       cat(paste(output, collapse = "\n"),"\n")
}
f <- data.frame(sim.data)
mod_stargazer(f, title="Simulated Data Homework 2")
```

Table 1: Simulated Data Homework 2

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
rural	35,000	0.700	0.458	0	0	1	1
armed	35,000	0.560	0.496	0	0	1	1

Linear Regression Model

```
model1 <-  glm(armed ~ rural, family=binomial(link="logit"), data=sim.data)
```

Regression table:

```
library("stargazer")

mod_stargazer <- function(...){
  output <- capture.output(stargazer(...))
  output <- output[4:length(output)]
  cat(paste(output, collapse = "\n"), "\n")
}

mod_stargazer(model1, type="latex",
          omit.table.layout="n",
          digits=2,
          digits.extra=0,
          title="Impact of Area of Living on Membership to an Armed Group")
```

Table 2: Impact of Area of Living on Membership to an Armed Group

	Dependent variable:
	armed
rural	1.26^{***}
	(0.02)
Constant	-0.64^{***}
	(0.02)
Observations	35,000
Log Likelihood	−22,613.09
Akaike Inf. Crit.	45,230.19