# Consillient Assessment

Chetan

11/03/2022

```
library(readxl)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(rstatix)
```

```
## Warning: package 'rstatix' was built under R version 4.2.2
```

```
##
## Attaching package: 'rstatix'
##
## The following object is masked from 'package:stats':
##
##     filter
```

```
data <- read_excel("./data.xlsx")
```

```
## New names:
## * `Please specify other` -> `Please specify other...7`
## * `Please specify other` -> `Please specify other...31`
## * `Please specify other` -> `Please specify other...33`
## * `Why not?` -> `Why not?...39`
## * `The most important thing in resolving the case is:` -> `The most important
##   thing in resolving the case is:...60`
## * `The second most important thing in resolving the case is:` -> `The second
##   most important thing in resolving the case is:...62`
## * `Please specify other` -> `Please specify other...64`
## * `Was this person in your diya group?` -> `Was this person in your diya
##   group?...66`
## * `Was this person in your family?` -> `Was this person in your family?...67`
## * `What was the marital status of the victim at the time of the incident?` ->
```

```
##   ‘What was the marital status of the victim at the time of the incident?...68‘
## * ‘Please specify other‘ -> ‘Please specify other...70‘
## * ‘As far as you know, who was the crime reported to next?‘ -> ‘As far as you
##   know, who was the crime reported to next?...71‘
## * ‘Please specify other‘ -> ‘Please specify other...72‘
## * ‘Was the incident ever reported to the police?‘ -> ‘Was the incident ever
##   reported to the police?...74‘
## * ‘Who ultimately resolved this case?‘ -> ‘Who ultimately resolved this
##   case?...75‘
## * ‘Please specify other‘ -> ‘Please specify other...76‘
## * ‘Which of the following were part of the final resolution? Please select all
##   that‘ -> ‘Which of the following were part of the final resolution? Please
##   select all that...77‘
## * ‘Please specify other‘ -> ‘Please specify other...85‘
## * ‘How long did it take to resolve the case?‘ -> ‘How long did it take to
##   resolve the case?...86‘
## * ‘Was the diya actually paid?‘ -> ‘Was the diya actually paid?...87‘
## * ‘How much was the diya payment? (PLEASE INCLUDE UNITS! Insert the amount that
##   was‘ -> ‘How much was the diya payment? (PLEASE INCLUDE UNITS! Insert the
##   amount that was...88‘
## * ‘Diya payments might go to the victim, the diya group, the elders, or others.
##   Do‘ -> ‘Diya payments might go to the victim, the diya group, the elders, or
##   others. Do...89‘
## * ‘Please specify other‘ -> ‘Please specify other...90‘
## * ‘As far as you know, did the victim feel that justice was served?‘ -> ‘As far
##   as you know, did the victim feel that justice was served?...91‘
## * ‘why not?‘ -> ‘why not?...92‘
## * ‘Why not?‘ -> ‘Why not?...94‘
## * ‘The most important thing in resolving the case is:‘ -> ‘The most important
##   thing in resolving the case is:...95‘
## * ‘The second most important thing in resolving the case is:‘ -> ‘The second
##   most important thing in resolving the case is:...97‘
## * ‘Please specify other‘ -> ‘Please specify other...99‘
## * ‘Was this person in your diya group?‘ -> ‘Was this person in your diya
##   group?...103‘
## * ‘Was this person in your family?‘ -> ‘Was this person in your family?...104‘
## * ‘What was the marital status of the victim at the time of the incident?‘ ->
##   ‘What was the marital status of the victim at the time of the
##   incident?...105‘
## * ‘Please specify other‘ -> ‘Please specify other...107‘
## * ‘As far as you know, who was the crime reported to next?‘ -> ‘As far as you
##   know, who was the crime reported to next?...108‘
## * ‘Please specify other‘ -> ‘Please specify other...110‘
## * ‘Was the incident ever reported to the police?‘ -> ‘Was the incident ever
##   reported to the police?...111‘
## * ‘Who ultimately resolved this case?‘ -> ‘Who ultimately resolved this
##   case?...112‘
## * ‘Please specify other‘ -> ‘Please specify other...113‘
## * ‘Which of the following were part of the final resolution? Please select all
##   that‘ -> ‘Which of the following were part of the final resolution? Please
##   select all that...114‘
## * ‘Please specify other‘ -> ‘Please specify other...122‘
## * ‘How long did it take to resolve the case?‘ -> ‘How long did it take to
##   resolve the case?...123‘
```

```
## * 'How much was the diya payment? (PLEASE INCLUDE UNITS! Insert the amount that
##   was' -> 'How much was the diya payment? (PLEASE INCLUDE UNITS! Insert the
##   amount that was...124'
## * 'Was the diya actually paid?' -> 'Was the diya actually paid?...125'
## * 'Diya payments might go to the victim, the diya group, the elders, or others.
##   Do' -> 'Diya payments might go to the victim, the diya group, the elders, or
##   others. Do...126'
## * 'As far as you know, did the victim feel that justice was served?' -> 'As far
##   as you know, did the victim feel that justice was served?...127'
## * 'why not?' -> 'why not?...128'
## * 'The most important thing in resolving the case is:' -> 'The most important
##   thing in resolving the case is:...129'
## * 'The second most important thing in resolving the case is:' -> 'The second
##   most important thing in resolving the case is:...131'
## * 'Please specify other' -> 'Please specify other...133'
```

## Descritive Statitics

```
##
##  0. No 1. Yes   <NA>
##    303    203    511
```

```
##
## 0. sgbv   1. dv    <NA>
##     506     511       0
```

```
#class(data_new$random_half)
#unique(data_new$random_half)

# Random half
data_new$random_half[which(data_new$random_half == '0. sgbv')] <- "sgbv"
data_new$random_half[which(data_new$random_half == '1. dv')] <- "dv"

# Female police
data_new$female_police[which(data_new$female_police == '0. No')] <- "No"
data_new$female_police[which(data_new$female_police == '1. Yes')] <- "Yes"
data_new$female_police[which(data_new$female_police == '98. Don\'t know')] <- "Don\'t know"

# Education
data_new$education[which(data_new$education == '6. College and above')] <- "College and above"
data_new$education[which(data_new$education == '0. None')] <- "None"
data_new$education[which(data_new$education == '1. Madrassa only')] <- "Madrassa"
data_new$education[which(data_new$education == '3. Completed Primary')] <- "Completed Primary"
data_new$education[which(data_new$education == '2. Some Primary')] <- "Some Primary"
data_new$education[which(data_new$education == '5. Completed Secondary')] <- "Completed Secondary"
data_new$education[which(data_new$education == '4. Some Secondary')] <- "Some Secondary"


# land own in
data_new$land_own_in[which(data_new$land_own_in == '1. Yes')] <- "Yes"
data_new$land_own_in[which(data_new$land_own_in == '0. No')] <- "No"

# marital status
```

```r
data_new$marital_status[which(data_new$marital_status == '1. Married')] <- "Married"
data_new$marital_status[which(data_new$marital_status == '2. Single')] <- "Single"
data_new$marital_status[which(data_new$marital_status == '3. Divorced')] <- "Divorced"
data_new$marital_status[which(data_new$marital_status == '5. Separated')] <- "Separated"
data_new$marital_status[which(data_new$marital_status == '4. Widow/widower')] <- "Widow/widower"
data_new$marital_status[which(data_new$marital_status == '99. Refused to answer')] <- "Refused to answe


# police loc
data_new$police_loc[which(data_new$police_loc == '1. Yes')] <- "Yes"
data_new$police_loc[which(data_new$police_loc == '0. No')] <- "No"

# time here
data_new$time_here[which(data_new$time_here == '3. More than 5 years')] <- "More than 5 years"
data_new$time_here[which(data_new$time_here == '2. 1-5 years')] <- "1-5 years"
data_new$time_here[which(data_new$time_here == '1. Less than one year')] <- "Less than one year"
data_new$time_here[which(data_new$time_here == '98. Don\'t know')] <- "Don\'t know"

## Walk Night
data_new$walk_night[which(data_new$walk_night == '2. Yes')] <- "Yes"
data_new$walk_night[which(data_new$walk_night == '0. No')] <- "No"
data_new$walk_night[which(data_new$walk_night == '1. Sometimes')] <- "Sometimes"

## Known victims
data_new$known_victim[which(data_new$known_victim == '1. Yes')] <- "Yes"
data_new$known_victim[which(data_new$known_victim == '0. No')] <- "No"

# sgbv_ideal_outcome
data_new$sgbv_ideal_outcome1_lab[which(data_new$sgbv_ideal_outcome1_lab == 'Dambiilaha waa la ciqaabay'
data_new$sgbv_ideal_outcome1_lab[which(data_new$sgbv_ideal_outcome1_lab == 'Qofka dhibanaha ahi magdhow
data_new$sgbv_ideal_outcome1_lab[which(data_new$sgbv_ideal_outcome1_lab == 'Qoyska/kooxda dhibbanaha wa

data_new <- data_new[-(which(data_new$sgbv_ideal_outcome1_lab == 'qofk dhibanaag magdhow yalasiiy mise



# dv_ideal_outcome1_lab
data_new$dv_ideal_outcome1_lab[which(data_new$dv_ideal_outcome1_lab == 'Qofka dhibanaha ahi magdhow aya
data_new$dv_ideal_outcome1_lab[which(data_new$dv_ideal_outcome1_lab == 'Dambiilaha waa la ciqaabay')] <
data_new$dv_ideal_outcome1_lab[which(data_new$dv_ideal_outcome1_lab == 'Qoyska/kooxda dhibbanaha waxaa

# Deleting na from random half
colSums(is.na(data_new))
```

```
##                      age          children_girls               education
##                        1                     162                       0
##          female_police                hh_size                     idp
##                        0                       0                       0
##             land_own_in               majority         marital_status
##                        0                      19                       0
##              police_loc                  rural              time_here
##                        0                       0                       0
##              walk_night           known_victim sgbv_ideal_outcome1_lab
```

```
##                       0                     511                     602
##   dv_ideal_outcome1_lab            random_half
##                     702                       0
```

```
unique(data_new$dv_ideal_outcome1_lab)
```

```
## [1] NA
## [2] "The individual victim is compensated or made whole or supported"
## [3] "The offender is punished"
## [4] "The victim's family / group is compensated or made whole or supported"
```

```
# Making factors
data_new$known_victim <- as.factor(data_new$known_victim)
data_new$walk_night <- as.factor(data_new$walk_night)
data_new$time_here <- factor(data_new$time_here, levels = c("Don't know","Less than one year","1-5 years

data_new$police_loc <- as.factor(data_new$police_loc)
data_new$marital_status <- as.factor(data_new$marital_status)
data_new$land_own_in <- as.factor(data_new$land_own_in)
data_new$education <- factor(data_new$education, levels = c("None", "Madrassa", "Some Primary", "Comple
data_new$female_police <- as.factor(data_new$female_police)
data_new$random_half <- as.factor(data_new$random_half)
data_new$idp <- as.factor(data_new$idp)
data_new$rural <- as.factor(data_new$rural)
data_new$majority <- as.factor(data_new$majority)
data_new$sgbv_ideal_outcome1_lab <- as.factor(data_new$sgbv_ideal_outcome1_lab)
data_new$dv_ideal_outcome1_lab <- as.factor(data_new$dv_ideal_outcome1_lab)
str(data_new)
```

```
## tibble [1,016 x 17] (S3: tbl_df/tbl/data.frame)
##  $ age                     : num [1:1016] NA 30 30 19 20 38 24 30 38 28 ...
##  $ children_girls          : num [1:1016] NA 4 5 2 NA 0 3 2 3 2 ...
##  $ education               : Ord.factor w/ 7 levels "None"<"Madrassa"<..: 7 1 2 4 4 3 2 2 1 1 ...
##  $ female_police           : Factor w/ 3 levels "Don't know","No",..: 2 3 1 2 3 2 2 2 2 2 ...
##  $ hh_size                 : num [1:1016] 7 12 10 5 12 4 6 10 10 9 ...
##  $ idp                     : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 2 2 1 1 ...
##  $ land_own_in             : Factor w/ 2 levels "No","Yes": 2 1 1 1 1 1 1 1 1 1 ...
##  $ majority                : Factor w/ 2 levels "majority","non majority": 2 2 1 2 1 2 2 2 2 2 ...
##  $ marital_status          : Factor w/ 6 levels "Divorced","Married",..: 2 2 2 2 5 1 4 4 2 2 ...
##  $ police_loc              : Factor w/ 2 levels "No","Yes": 2 2 2 1 2 2 2 2 2 2 ...
##  $ rural                   : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 1 1 1 ...
##  $ time_here               : Ord.factor w/ 4 levels "Don't know"<"Less than one year"<..: 4 3 4 4 4 4
##  $ walk_night              : Factor w/ 3 levels "No","Sometimes",..: 1 3 3 3 1 3 3 1 1 3 ...
##  $ known_victim            : Factor w/ 2 levels "No","Yes": 2 NA 1 2 2 NA NA 1 NA NA ...
##  $ sgbv_ideal_outcome1_lab: Factor w/ 3 levels "The individual victim is compensated or made whole o
##  $ dv_ideal_outcome1_lab  : Factor w/ 3 levels "The individual victim is compensated or made whole o
##  $ random_half             : Factor w/ 2 levels "dv","sgbv": 2 1 2 2 2 1 1 2 1 1 ...
```

```
#class(data_new$time_here)

data.group.random.half <- group_by(data_new, random_half)
#get_summary_stats(data.group.random.half, children_girls, type = "mean_sd")
```

```
table(data.group.random.half$random_half, data.group.random.half$female_police)
```

```
##
##        Don't know  No Yes
##   dv           55 249 207
##   sgbv         60 255 190
```

```
summary(data_new)
```

```
##       age        children_girls                  education      female_police
##  Min.   :15.00   Min.   : 0.000   None               :473   Don't know:115
##  1st Qu.:25.00   1st Qu.: 1.000   Madrassa           :241   No         :504
##  Median :30.00   Median : 3.000   Some Primary       :128   Yes        :397
##  Mean   :32.62   Mean   : 2.738   Completed Primary  : 56
##  3rd Qu.:39.00   3rd Qu.: 4.000   Some Secondary     : 29
##  Max.   :88.00   Max.   :10.000   Completed Secondary: 42
##  NA's   :1       NA's   :162      College and above  : 47
##     hh_size       idp     land_own_in         majority
##  Min.   : 0.000   0:804   No :765     majority    :307
##  1st Qu.: 6.000   1:212   Yes:251     non majority:690
##  Median : 9.000                       NA's        : 19
##  Mean   : 8.885
##  3rd Qu.:11.000
##  Max.   :30.000
##
##            marital_status police_loc rural                time_here
##  Divorced         :156    No :195    0:910   Don't know        :  6
##  Married          :635    Yes:821    1:106   Less than one year: 22
##  Refused to answer:  1                       1-5 years         :233
##  Separated        : 14                       More than 5 years :755
##  Single           :131
##  Widow/widower    : 79
##
##      walk_night  known_victim
##  No       :409   No  :302
##  Sometimes: 77   Yes :203
##  Yes      :530   NA's:511
##
##
##
##
##                                                              sgbv_ideal_outcome1_lab
##  The individual victim is compensated or made whole or supported      :117
##  The offender is punished                                             :228
##  The victim's family / group is compensated or made whole or supported: 69
##  NA's                                                                 :602
##
##
##
##                                                               dv_ideal_outcome1_lab
##  The individual victim is compensated or made whole or supported      :109
##  The offender is punished                                             :141
```

```
##  The victim's family / group is compensated or made whole or supported: 64
##  NA's                                                                   :702
##
##
##
##  random_half
##  dv  :511
##  sgbv:505
##
##
##
##
##
```

```
data_sgbv <- filter(data_new, random_half == 'sgbv')

#table(data.group.random.half$random_half, data.group.random.half$known_victim)
# table(data.group.random.half$random_half, data.group.random.half$female_police)
# table(data.group.random.half$random_half, data.group.random.half$time_here)
```

```
#ggplot(data_sgbv, aes(x = walk_night, y = age)) + geom_bar(aes(fill = majority), stat = "identity", co

# table_sgbv_rural_time <- table(data_sgbv$rural, data_sgbv$time_here)
#
# addmargins(table_sgbv_rural_time)
# addmargins(prop.table(table_sgbv_rural_time))


table_sgbv_rural_police <- table(data_sgbv$rural, data_sgbv$police_loc)

addmargins(table_sgbv_rural_police)
```

```
##
##        No Yes Sum
##   0    71 377 448
##   1    17  40  57
##   Sum  88 417 505
```

```
addmargins(prop.table(table_sgbv_rural_police))
```

```
##
##               No        Yes        Sum
##   0   0.14059406 0.74653465 0.88712871
##   1   0.03366337 0.07920792 0.11287129
##   Sum 0.17425743 0.82574257 1.00000000
```

```
chisq.test(data_sgbv$rural, data_sgbv$police_loc)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  data_sgbv$rural and data_sgbv$police_loc
## X-squared = 5.9276, df = 1, p-value = 0.01491
```

Based on the results of the chi-square test generated above, p-value is 0.014, which is less than 0.05. This implies that there is a statistically significant relationship between the rural and police_loc variable.

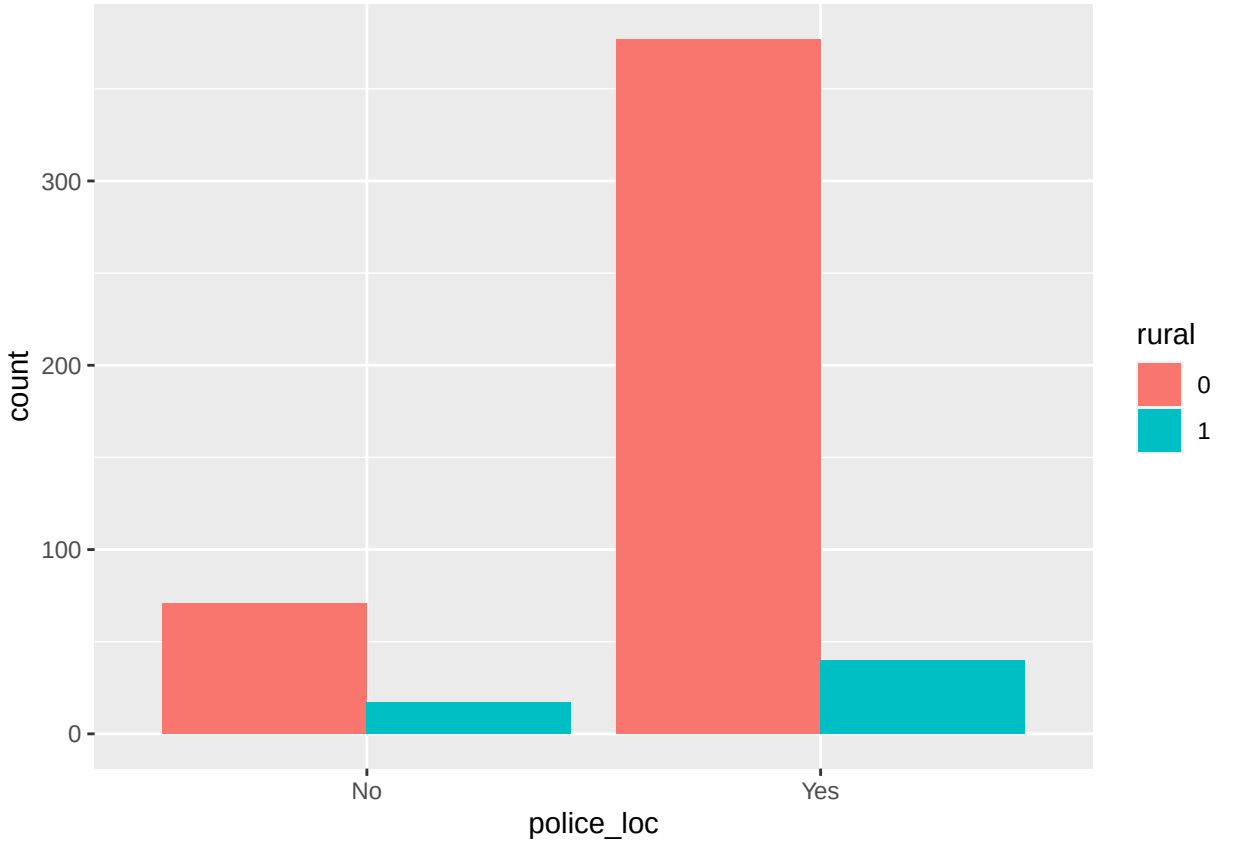
```
# sgbv_table <- data_sgbv %>% count(rural, time_here, sort = TRUE)
# sgbv_table

sgbv_table <- data_sgbv %>% count(police_loc, rural, sort = TRUE)
sgbv_table
```

```
## # A tibble: 4 x 3
##   police_loc rural     n
##   <fct>      <fct> <int>
## 1 Yes        0       377
## 2 No         0        71
## 3 Yes        1        40
## 4 No         1        17
```

```
# ggplot(data = data_sgbv) +
#   geom_bar(mapping = aes(x = time_here,  fill = rural), position = "dodge")

ggplot(data = data_sgbv) +
  geom_bar(mapping = aes(x = police_loc,  fill = rural), position = "dodge")
```




From the above bar graph it is evident that people who have stayed more than 5 years in urban area know more about sgbv cases. Also, people who live in urban areas and has a police station nearby also knows more about such cases.

### Question 2:

```
(ideal_random_table <- data_new %>% count(sgbv_ideal_outcome1_lab, dv_ideal_outcome1_lab, random_half, s
```

```
## # A tibble: 8 x 4
##   sgbv_ideal_outcome1_lab                                    dv_id~1 rando~2     n
##   <fct>                                                      <fct>   <fct>   <int>
## 1 The offender is punished                                   <NA>    sgbv      228
## 2 <NA>                                                       <NA>    dv        197
## 3 <NA>                                                       The of~ dv        141
## 4 The individual victim is compensated or made whole or s~ <NA>    sgbv      117
## 5 <NA>                                                       The in~ dv        109
## 6 <NA>                                                       <NA>    sgbv       91
## 7 The victim's family / group is compensated or made whol~ <NA>    sgbv       69
## 8 <NA>                                                       The vi~ dv         64
## # ... with abbreviated variable names 1: dv_ideal_outcome1_lab, 2: random_half
```

```
ideal_random_table_new <- ideal_random_table[!with(ideal_random_table,is.na(dv_ideal_outcome1_lab)& is.
```

```
for(i in 1:length(ideal_random_table_new$sgbv_ideal_outcome1_lab)){

    if(!is.na(ideal_random_table_new$sgbv_ideal_outcome1_lab[i])){
        ideal_random_table_new$dv_ideal_outcome1_lab[i] <-  ideal_random_table_new$sgbv_ideal_outcome1_l
  }
}

ideal_random_table_new <- ideal_random_table_new[,-1]
names(ideal_random_table_new) <-  c("ideal_outcome1_lab","random_half","n")
names(ideal_random_table_new)
```

```
## [1] "ideal_outcome1_lab" "random_half"        "n"
```

```
#length(ideal_random_table_new)
#require(cowplot)
# (plot1 <- ggplot(data = data_new) +
#   stat_count(mapping = aes(x = dv_ideal_outcome1_lab)))
# (plot2 <- ggplot(data = data_new) +
#   stat_count(mapping = aes(x = sgbv_ideal_outcome1_lab)))
#plot_grid(plot1, plot2, labels = "AUTO")


ideal_random_table_new
```

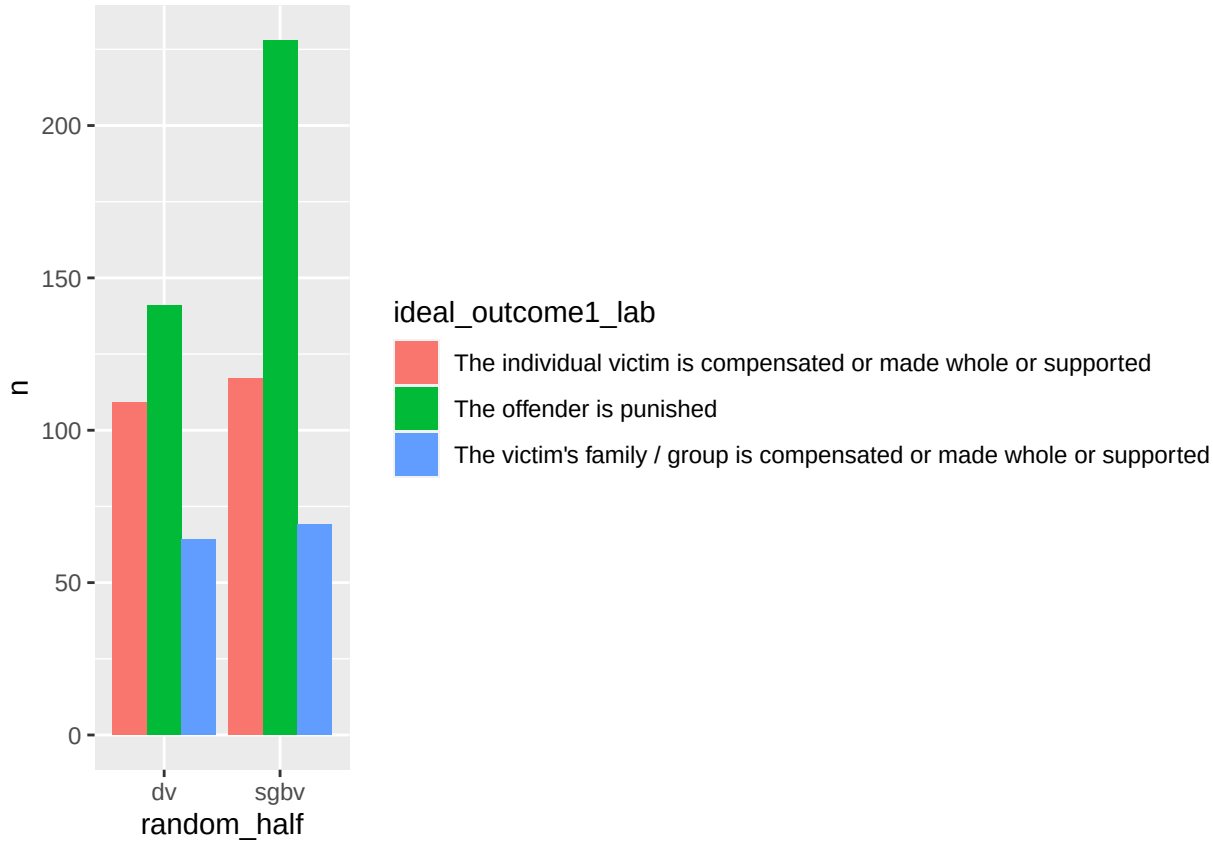
```
## # A tibble: 6 x 3
##   ideal_outcome1_lab                                                rando~1     n
##   <fct>                                                             <fct>   <int>
## 1 The offender is punished                                          sgbv      228
## 2 The offender is punished                                          dv        141
## 3 The individual victim is compensated or made whole or supported   sgbv      117
## 4 The individual victim is compensated or made whole or supported   dv        109
```

```
## 5 The victim's family / group is compensated or made whole or sup~ sgbv       69
## 6 The victim's family / group is compensated or made whole or sup~ dv         64
## # ... with abbreviated variable name 1: random_half
```

```
str(ideal_random_table_new)
```

```
## tibble [6 x 3] (S3: tbl_df/tbl/data.frame)
##  $ ideal_outcome1_lab: Factor w/ 3 levels "The individual victim is compensated or made whole or supp
##  $ random_half       : Factor w/ 2 levels "dv","sgbv": 2 1 2 1 2 1
##  $ n                 : int [1:6] 228 141 117 109 69 64
```

```
ggplot(data = ideal_random_table_new) +
  geom_bar(mapping = aes(x = random_half,y = n, fill = ideal_outcome1_lab), position = "dodge", stat =
```

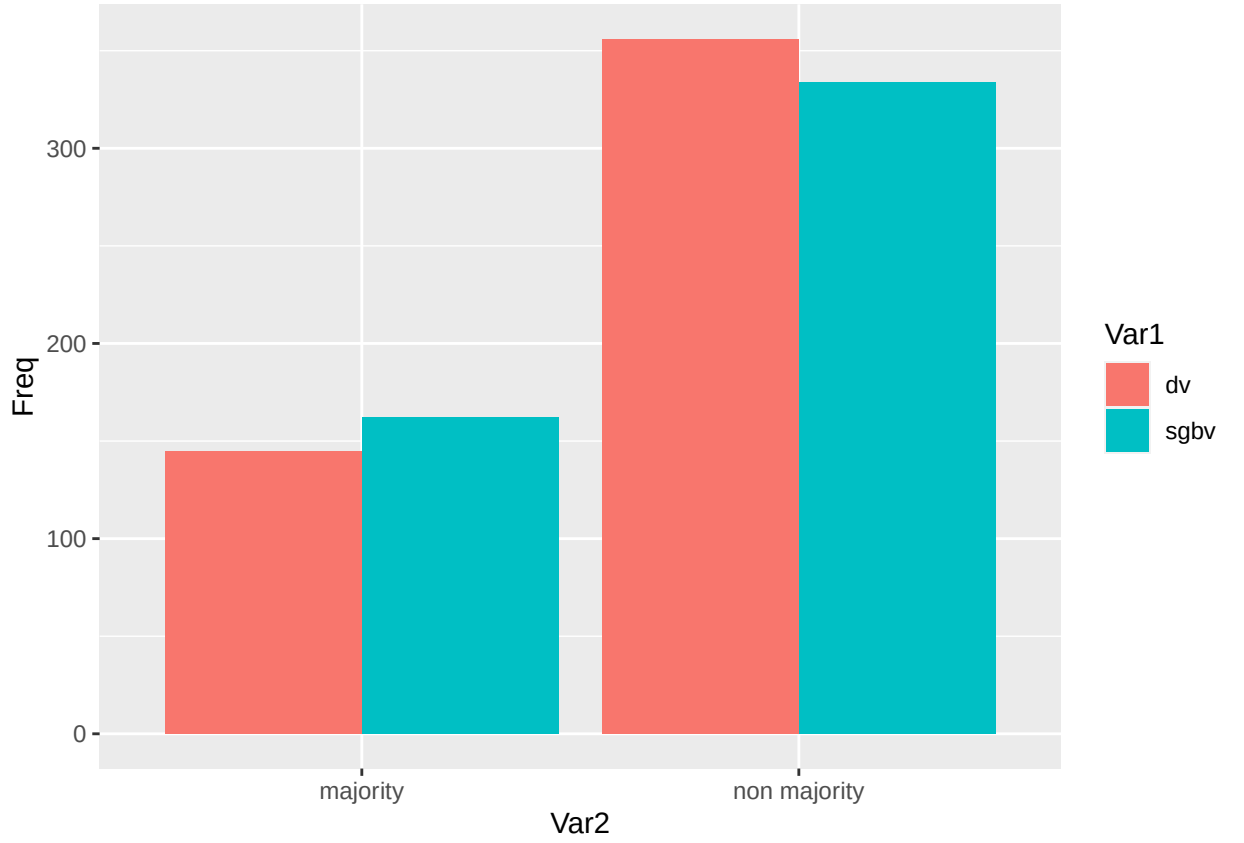



## Answer 3

```
(data_majority_table <- data.frame(table(data.group.random.half$random_half, data.group.random.half$majo
```

```
##   Var1         Var2 Freq
## 1   dv     majority  145
## 2 sgbv     majority  162
## 3   dv non majority  356
## 4 sgbv non majority  334
```

```
ggplot(data = data_majority_table) +
  geom_bar(mapping = aes(x = Var2 ,y = Freq, fill = Var1), position = "dodge", stat = "identity")
```




From the above bar graph it, we can see non-majority people in the clan are more tend to dv and sgbv then majority people.