# Project Proposal

**Due November 17 at 11:59pm**

Afag, Peter, Elenor, Mobasser

### Load Packages

```
library(tidyverse)
library(dplyr)
library(ggplot2)
library(cowplot)
```

## Dataset 1 (top choice)

Police Killing Dataset

**Data source:** https://github.com/fivethirtyeight/data/tree/master/police-killings

**Brief description:** This dataset includes details on police killings in the United States, with variables such as location, race, gender, age, and circumstances of each incident.

**Research question 1: Is there an association between the economic conditions of a county and the racial composition of individuals involved in police-related incidents?**

- Outcome variable: race ethnicity (Nominal) – categorizes the racial/ethnic group of the deceased.
- Independent Variables:
  - *compincome* : A measure of relative income (household income divided by county income) which indicates economic status.
  - *pov* : Poverty level in the area, providing additional context on socio-economic conditions.

**Research question 2: Does the likelihood of an individual being armed in a police-related incident vary based on age and unemployment rate in the area?**

- Outcome variable: armed (Binary) – indicates whether the deceased was armed at the time of the incident.
- Independent Variables:
  - *age*: Age of the deceased (continuous variable).
  - *urate*: Unemployment rate in the area, which could represent socio-economic stress.
- Interaction Term: An interaction between age and urate to see if age and local unemployment influence the likelihood of individuals being armed.

**Load the data and provide a `glimpse()`:**

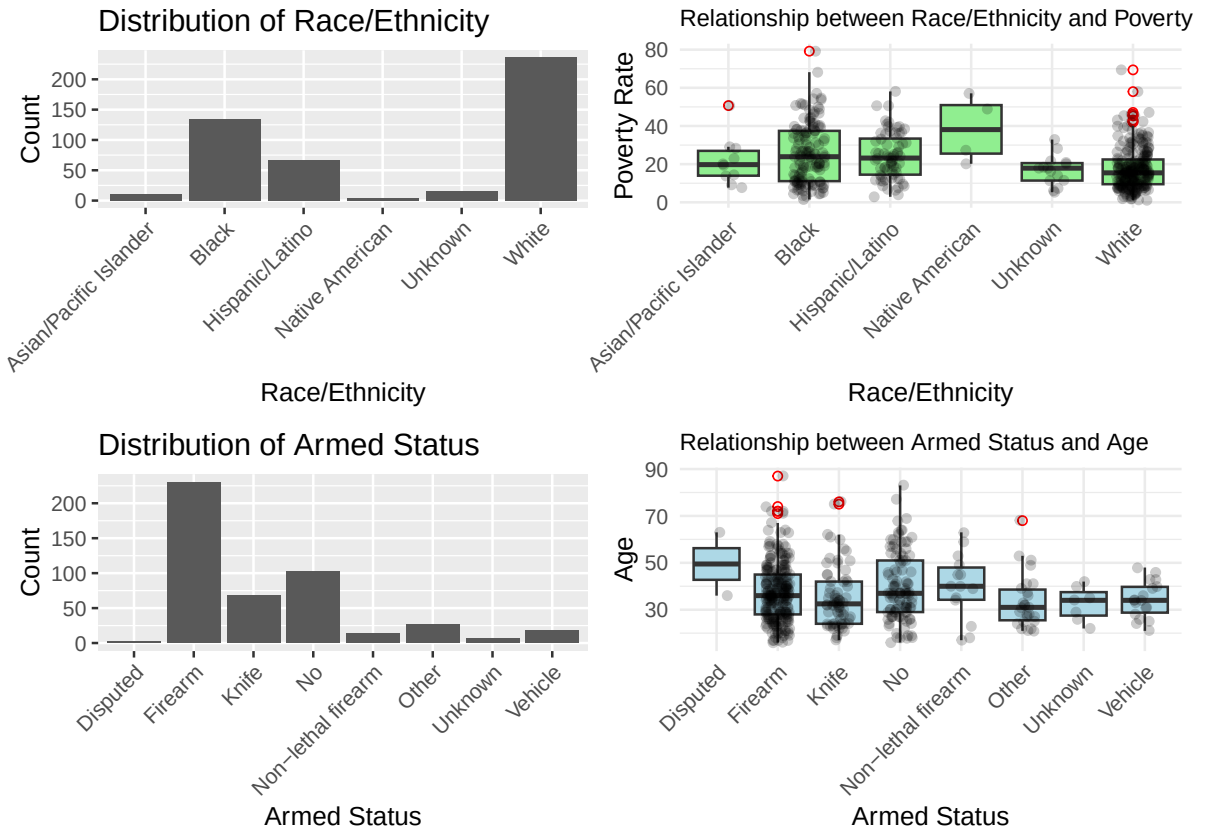
```
new_data <-  read.csv("data/police_killings.csv")
glimpse(new_data)
```

```
Rows: 467
Columns: 34
$ name                 <chr> "A'donte Washington", "Aaron Rutledge", "Aaron Si~
$ age                  <chr> "16", "27", "26", "25", "29", "29", "22", "35", "~
$ gender               <chr> "Male", "Male", "Male", "Male", "Male", "Male", "~
$ raceethnicity        <chr> "Black", "White", "White", "Hispanic/Latino", "Wh~
$ month                <chr> "February", "April", "March", "March", "March", "~
$ day                  <int> 23, 2, 14, 11, 19, 7, 27, 26, 28, 7, 26, 12, 20, ~
$ year                 <int> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2~
$ streetaddress        <chr> "Clearview Ln", "300 block Iris Park Dr", "22nd A~
$ city                 <chr> "Millbrook", "Pineville", "Kenosha", "South Gate"~
$ state                <chr> "AL", "LA", "WI", "CA", "OH", "AZ", "CA", "CA", "~
$ latitude             <dbl> 32.52958, 31.32174, 42.58356, 33.93930, 41.14857,~
$ longitude            <dbl> -86.36283, -92.43486, -87.83571, -118.21946, -81.~
$ state_fp             <int> 1, 22, 55, 6, 39, 4, 6, 6, 48, 26, 6, 6, 48, 18, ~
$ county_fp            <int> 51, 79, 59, 37, 153, 13, 29, 37, 41, 81, 31, 59, ~
$ tract_ce             <int> 30902, 11700, 1200, 535607, 530800, 111602, 700, ~
$ geo_id               <dbl> 1051030902, 22079011700, 55059001200, 6037535607,~
$ county_id            <int> 1051, 22079, 55059, 6037, 39153, 4013, 6029, 6037~
$ namelsad             <chr> "Census Tract 309.02", "Census Tract 117", "Censu~
$ lawenforcementagency <chr> "Millbrook Police Department", "Rapides Parish Sh~
$ cause                <chr> "Gunshot", "Gunshot", "Gunshot", "Gunshot", "Guns~
$ armed                <chr> "No", "No", "No", "Firearm", "No", "No", "Firearm~
$ pop                  <int> 3779, 2769, 4079, 4343, 6809, 4682, 5027, 5238, 4~
$ share_white          <chr> "60.5", "53.8", "73.8", "1.2", "92.5", "7", "50.8~
$ share_black          <chr> "30.5", "36.2", "7.7", "0.6", "1.4", "7.7", "0.3"~
$ share_hispanic       <chr> "5.6", "0.5", "16.8", "98.8", "1.7", "79", "44.2"~
$ p_income             <chr> "28375", "14678", "25286", "17194", "33954", "155~
```

```
$ h_income             <int> 51367, 27972, 45365, 48295, 68785, 20833, 58068, ~
$ county_income        <int> 54766, 40930, 54930, 55909, 49669, 53596, 48552, ~
$ comp_income          <dbl> 0.9379359, 0.6834107, 0.8258693, 0.8638144, 1.384~
$ county_bucket        <int> 3, 2, 2, 3, 5, 1, 4, 4, 2, 3, 4, 5, 3, 4, 3, 1, 3~
$ nat_bucket           <int> 3, 1, 3, 3, 4, 1, 4, 4, 1, 2, 3, 5, 3, 2, 2, 1, 3~
$ pov                  <chr> "14.1", "28.8", "14.6", "11.7", "1.9", "58", "17.~
$ urate                <dbl> 0.09768638, 0.06572379, 0.16629314, 0.12482727, 0~
$ college              <dbl> 0.16850951, 0.11140236, 0.14731227, 0.05013293, 0~
```

```r
# Convert "Unknown" age to mean age
new_data$age <- as.numeric(ifelse(new_data$age == "unknown", NA, new_data$age))
new_data$age[is.na(new_data$age)] <- mean(new_data$age, na.rm = TRUE)
# convert "-" pov rate to median pov
new_data$pov <- as.numeric(ifelse(new_data$pov == "-", NA, new_data$pov))
new_data$pov[is.na(new_data$pov)] <- median(new_data$pov, na.rm = TRUE)

####### Research question 1
## 1. Race/Ethnicity

plot1 <- ggplot(new_data, aes(x = raceethnicity)) +
  geom_bar() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Distribution of Race/Ethnicity", x = "Race/Ethnicity", y = "Count")

# Plot 1: Relationship between 'race' and 'poverty rate'
plot2 <- ggplot(new_data, aes(x = raceethnicity, y = pov)) +
  geom_boxplot(fill = "lightgreen", outlier.colour = "red", outlier.shape = 1) +
  geom_jitter(width = 0.2, alpha = 0.2) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_text(size = 10)) +
  labs(title = "Relationship between Race/Ethnicity and Poverty Rate",
       x = "Race/Ethnicity",
       y = "Poverty Rate")

#### Research question 2

## 6. Armed Status

plot3 <- ggplot(new_data, aes(x = armed)) +
  geom_bar() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
```

```
  labs(title = "Distribution of Armed Status", x = "Armed Status", y = "Count")

# Plot 2: Relationship between 'armed' and 'age'
plot4 <- ggplot(new_data, aes(x = armed, y = age)) +
  geom_boxplot(fill = "lightblue", outlier.colour = "red", outlier.shape = 1) +
  geom_jitter(width = 0.2, alpha = 0.2) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_text(size = 10)) +
  labs(title = "Relationship between Armed Status and Age",
       x = "Armed Status",
       y = "Age")

plot_grid(plot1, plot2, plot3, plot4, ncol=2)
```




There are mostly black and white in race/ethnicity in police killings. We can see that the average poverty rate is lower in white compared to black. Moreover, we can observe that there are many people armed with firearm and the range of the age of the deceased is wide from less than 30 to 90.

## Dataset 2

**Data source:** Pew Research East Asian Religion

**Brief description:** The dataset surveys different east asian countries to see how many of them are religious and how the factors that influence them

**Research question 1:**

What variables are associated with being religious?

- Outcome variable: Religion importance - Q9. Ordinal categorical variable, the lower it is the more important religion is to the survey taker.

**Research question 2:**

Do different regions have an effect on the factors that correlates with a person's bent on religion?

- Outcome variable (include the name/description and type of variable): Religion importance - Q9. Ordinal categorical variable, the lower it is the more important religion is to the survey taker.

**Load the data and provide a `glimpse()`:**

```
library(tidyverse)

data <- read.csv("data/dataset_2.csv")

set.seed(12)
data_subset <- data %>% sample_n(500) %>% filter(Q9 < 5)

summary(data_subset)
```

```
      QRID              weight        SurveyPublic      DEMONYM
 Min.   : 1300002   Min.   :0.2287   Min.   :1.000   Min.   :1.000
 1st Qu.: 4039246   1st Qu.:0.5071   1st Qu.:2.000   1st Qu.:2.000
 Median : 5502516   Median :0.8039   Median :4.000   Median :4.000
 Mean   : 7397161   Mean   :0.9943   Mean   :3.694   Mean   :3.694
 3rd Qu.: 6009246   3rd Qu.:1.2039   3rd Qu.:5.000   3rd Qu.:5.000
 Max.   :23255194   Max.   :3.6494   Max.   :6.000   Max.   :6.000

       Q1              Q2a             Q2b             Q2c
 Min.   : 1.000   Min.   :1.000   Min.   : 1.000   Min.   : 1.0
 1st Qu.: 1.000   1st Qu.:3.000   1st Qu.: 3.000   1st Qu.: 3.0
```

```
Median : 2.000   Median :3.000   Median : 3.000   Median : 3.0
Mean   : 4.069   Mean   :2.711   Mean   : 3.109   Mean   : 3.2
3rd Qu.: 3.000   3rd Qu.:3.000   3rd Qu.: 3.000   3rd Qu.: 3.0
Max.   :98.000   Max.   :4.000   Max.   :98.000   Max.   :98.0

     Q2d              Q3              Q4            QLEGAL
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 3.000   1st Qu.: 2.000   1st Qu.: 3.000   1st Qu.: 2.000
Median : 3.000   Median : 2.000   Median : 3.000   Median : 2.000
Mean   : 3.142   Mean   : 2.848   Mean   : 4.006   Mean   : 7.079
3rd Qu.: 3.000   3rd Qu.: 2.000   3rd Qu.: 3.000   3rd Qu.: 3.000
Max.   :98.000   Max.   :99.000   Max.   :98.000   Max.   :99.000

     Q5a              Q5b              Q5c              Q5d
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.00
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.00
Median : 2.000   Median : 1.000   Median : 2.000   Median : 1.00
Mean   : 4.244   Mean   : 3.393   Mean   : 3.751   Mean   : 2.66
3rd Qu.: 3.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.00
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.00
NA's   :100      NA's   :100      NA's   :100      NA's   :100
   QCURREL          Q6a              Q6b              Q6c
Min.   : 1.00   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.00   1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 1.000
Median : 7.50   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 5.63   Mean   : 2.559   Mean   : 3.164   Mean   : 3.915
3rd Qu.: 8.00   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :99.00   Max.   :98.000   Max.   :98.000   Max.   :98.000
                                                  NA's   :412
     Q6d              Q6e              Q6f            QBORN
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 2.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 7.415   Mean   : 7.134   Mean   : 3.986   Mean   : 6.938
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :99.000   Max.   :99.000
                                                  NA's   :382
     Q7a              Q7b              Q7c              Q7d
Min.   : 1.00   Min.   : 1.000   Min.   : 1.00   Min.   : 1.000
1st Qu.: 2.00   1st Qu.: 2.000   1st Qu.: 1.00   1st Qu.: 1.000
Median : 2.00   Median : 2.000   Median : 1.00   Median : 1.000
Mean   : 5.64   Mean   : 9.152   Mean   : 7.65   Mean   : 7.955
3rd Qu.: 2.00   3rd Qu.: 2.000   3rd Qu.: 2.00   3rd Qu.: 2.000
```

```
Max.   :99.00   Max.   :99.000   Max.   :99.00   Max.   :99.000

     Q7e              Q8a              Q8b              Q8c
Min.   : 1.000   Min.   : 1.00   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 2.00   1st Qu.: 1.000   1st Qu.: 2.000
Median : 1.000   Median : 2.00   Median : 2.000   Median : 2.000
Mean   : 5.684   Mean   : 8.94   Mean   : 9.534   Mean   : 9.699
3rd Qu.: 2.000   3rd Qu.: 2.00   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :99.000   Max.   :98.00   Max.   :98.000   Max.   :98.000
                 NA's   :361     NA's   :361      NA's   :361
     Q8d              Q8e              Q8f              Q8g
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000
Median : 1.000   Median : 2.000   Median : 1.000   Median : 1.000
Mean   : 5.797   Mean   : 6.587   Mean   : 7.218   Mean   : 8.541
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.000
NA's   :361      NA's   :361      NA's   :361      NA's   :361
     Q8h              Q8i              Q8j              Q8k
Min.   : 1.00   Min.   : 1.000   Min.   : 1.00   Min.   : 1.000
1st Qu.: 1.00   1st Qu.: 1.000   1st Qu.: 1.00   1st Qu.: 1.000
Median : 2.00   Median : 1.000   Median : 2.00   Median : 2.000
Mean   :12.46   Mean   : 5.526   Mean   :12.44   Mean   : 6.579
3rd Qu.: 2.00   3rd Qu.: 1.000   3rd Qu.: 2.00   3rd Qu.: 2.000
Max.   :98.00   Max.   :98.000   Max.   :98.00   Max.   :98.000
NA's   :361     NA's   :361      NA's   :361     NA's   :361
      Q9              Q10              Q11              Q12
Min.   :1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.:2.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 2.000
Median :3.000   Median : 1.000   Median : 1.000   Median : 2.000
Mean   :2.593   Mean   : 4.211   Mean   : 3.267   Mean   : 6.996
3rd Qu.:3.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :4.000   Max.   :99.000   Max.   :98.000   Max.   :99.000

    QCHREL           Q13              Q14              Q15a
Min.   : 1.000   Min.   : 1.000   Min.   :1.00   Min.   : 1.00
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.:1.00   1st Qu.: 1.00
Median : 5.000   Median : 1.000   Median :2.00   Median : 2.00
Mean   : 6.287   Mean   : 3.509   Mean   :2.03   Mean   : 2.02
3rd Qu.: 8.000   3rd Qu.: 3.000   3rd Qu.:3.00   3rd Qu.: 2.00
Max.   :98.000   Max.   :99.000   Max.   :3.00   Max.   :98.00
                 NA's   :384      NA's   :395
     Q15b             Q15c             Q15d          QFIGURESa
```

```
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 2.000   1st Qu.: 2.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 1.743   Mean   : 2.427   Mean   : 2.682   Mean   : 3.061
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.000

  QFIGURESb        QFIGURESc        QFIGURESd        QFIGURESe
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 2.000   1st Qu.: 2.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 2.198   Mean   : 2.097   Mean   : 5.891   Mean   : 1.998
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :99.000   Max.   :98.000

  QFIGURESf        QFIGURESg        QFIGURESh        QFIGURESi
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 2.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 2.148   Mean   : 2.595   Mean   : 2.768   Mean   : 3.688
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.000
                 NA's   :185      NA's   :412
    Q16_1            Q16_2            Q16_3            Q16_98           Q16_99
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :2
1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2
Median :1.000   Median :1.000   Median :2.000   Median :2.000   Median :2
Mean   :1.273   Mean   :1.437   Mean   :1.978   Mean   :1.994   Mean   :2
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2
Max.   :2.000   Max.   :2.000   Max.   :2.000   Max.   :2.000   Max.   :2

     Q17              Q18              Q19              Q20a
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000
Median : 2.000   Median : 2.000   Median : 1.000   Median : 1.000
Mean   : 3.028   Mean   : 2.879   Mean   : 2.038   Mean   : 4.245
3rd Qu.: 2.000   3rd Qu.: 3.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.000
NA's   :135                                        NA's   :225
     Q20b             Q21              Q22a             Q22b
Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.: 1.000
Median : 1.000   Median :3.000   Median :1.000   Median : 1.000
```

```
Mean   : 1.446   Mean   :2.673   Mean   :1.304   Mean   : 1.551
3rd Qu.: 1.000   3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :5.000   Max.   :2.000   Max.   :98.000
NA's   :225      NA's   :225
     Q22c            Q22d            Q23             Q24a
Min.   :1.000   Min.   :1.00   Min.   : 1.00   Min.   : 1.000
1st Qu.:1.000   1st Qu.:1.00   1st Qu.: 1.00   1st Qu.: 1.000
Median :1.000   Median :2.00   Median : 2.00   Median : 2.000
Mean   :1.267   Mean   :1.54   Mean   : 1.87   Mean   : 4.328
3rd Qu.:2.000   3rd Qu.:2.00   3rd Qu.: 2.00   3rd Qu.: 2.000
Max.   :2.000   Max.   :2.00   Max.   :98.00   Max.   :98.000

     Q24b            Q24c            Q24d            Q25a
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 1.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 2.000
Mean   : 4.439   Mean   : 2.506   Mean   : 2.429   Mean   : 3.502
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :98.000

     Q25b            Q25c            Q26a            Q26b
Min.   : 1.000   Min.   : 1.000   Min.   : 1.00   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 3.00   1st Qu.: 2.000
Median : 2.000   Median : 2.000   Median : 4.00   Median : 3.000
Mean   : 4.706   Mean   : 3.885   Mean   : 4.86   Mean   : 3.911
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 5.00   3rd Qu.: 4.000
Max.   :98.000   Max.   :99.000   Max.   :99.00   Max.   :98.000

     Q27a            Q27b            Q27c            Q27d
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.00
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.00
Median : 2.000   Median : 1.000   Median : 1.000   Median : 2.00
Mean   : 5.583   Mean   : 3.233   Mean   : 7.387   Mean   :21.13
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.00
Max.   :98.000   Max.   :99.000   Max.   :98.000   Max.   :99.00

     Q27e            Q27f            Q27g            Q27h
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 1.000
Median : 1.000   Median : 2.000   Median : 1.000   Median : 2.000
Mean   : 5.549   Mean   : 4.431   Mean   : 2.522   Mean   : 7.168
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :98.000   Max.   :99.000
```

```
     Q27i            Q28             Q29             Q30
Min.   : 1.00   Min.   : 1.00   Min.   : 1.000   Min.   : 1.00
1st Qu.: 1.00   1st Qu.: 5.00   1st Qu.: 1.000   1st Qu.: 1.00
Median : 1.00   Median : 6.00   Median : 2.000   Median : 2.00
Mean   : 6.34   Mean   : 5.66   Mean   : 1.763   Mean   : 4.15
3rd Qu.: 2.00   3rd Qu.: 7.00   3rd Qu.: 2.000   3rd Qu.: 2.00
Max.   :99.00   Max.   :98.00   Max.   :98.000   Max.   :99.00

     Q31             Q33             Q35             Q36
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 1.000
Median : 2.000   Median : 2.000   Median : 2.000   Median : 1.000
Mean   : 2.437   Mean   : 1.998   Mean   : 4.237   Mean   : 2.666
3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
Max.   :98.000   Max.   :98.000   Max.   :99.000   Max.   :98.000

     Q37             Q38a            Q38b            Q38c
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.00
1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 1.00
Median : 1.000   Median : 2.000   Median : 2.000   Median : 2.00
Mean   : 2.662   Mean   : 7.091   Mean   : 8.399   Mean   : 8.31
3rd Qu.: 1.000   3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.: 2.00
Max.   :99.000   Max.   :99.000   Max.   :99.000   Max.   :99.00
NA's   :100                                       NA's   :100
     Q39a            Q39b            Q39c            Q39d
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 3.000
Median : 2.000   Median : 3.000   Median : 2.000   Median : 3.000
Mean   : 5.164   Mean   : 7.014   Mean   : 3.911   Mean   : 6.468
3rd Qu.: 2.000   3rd Qu.: 4.000   3rd Qu.: 2.000   3rd Qu.: 4.000
Max.   :99.000   Max.   :99.000   Max.   :98.000   Max.   :99.000

    QABORT           Q40a            Q40b            Q40c
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   : 1.000
1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.: 3.000   1st Qu.: 3.000
Median : 2.000   Median : 2.000   Median : 4.000   Median : 4.000
Mean   : 7.006   Mean   : 3.459   Mean   : 4.982   Mean   : 5.343
3rd Qu.: 3.000   3rd Qu.: 2.000   3rd Qu.: 4.000   3rd Qu.: 4.000
Max.   :99.000   Max.   :98.000   Max.   :99.000   Max.   :99.000
                 NA's   :100      NA's   :100      NA's   :100
     Q40d            Q40e          QATTEND           QAGE
Min.   : 1.00   Min.   : 1.000   Min.   : 1.000   Min.   : 18.00
```

```
1st Qu.: 2.00   1st Qu.: 3.250   1st Qu.: 3.000   1st Qu.: 32.00
Median : 3.00   Median : 4.000   Median : 3.000   Median : 45.00
Mean   : 5.32   Mean   : 7.264   Mean   : 5.234   Mean   : 77.21
3rd Qu.: 4.00   3rd Qu.: 4.000   3rd Qu.: 5.000   3rd Qu.: 61.00
Max.   :99.00   Max.   :99.000   Max.   :98.000   Max.   :999.00
NA's   :100     NA's   :100      NA's   :430
  QAGE_NRHKG      QAGE_NRJPN      QAGE_NRKOR  QAGE_NRTWNrec   QAGE_NRVNM
Min.   :2.000   Min.   : 3.00   Min.   :2     Min.   :2.0   Min.   : NA
1st Qu.:3.000   1st Qu.: 3.50   1st Qu.:2     1st Qu.:2.5   1st Qu.: NA
Median :4.000   Median : 4.00   Median :2     Median :3.0   Median : NA
Mean   :3.778   Mean   :35.33   Mean   :2     Mean   :3.0   Mean   :NaN
3rd Qu.:5.000   3rd Qu.:51.50   3rd Qu.:2     3rd Qu.:3.5   3rd Qu.: NA
Max.   :5.000   Max.   :99.00   Max.   :2     Max.   :4.0   Max.   : NA
NA's   :485     NA's   :491     NA's   :493   NA's   :491   NA's   :494
   QETHHKG         QETHTWN         QETHVNM          QFERT            QHH1
Min.   : 1.00   Min.   : 1.00   Min.   : 1.00   Min.   : 0.00   Min.   :1.00
1st Qu.: 1.00   1st Qu.: 1.00   1st Qu.: 1.00   1st Qu.: 0.00   1st Qu.:2.00
Median : 1.00   Median : 1.00   Median : 1.00   Median : 2.00   Median :4.00
Mean   : 2.96   Mean   :15.96   Mean   :12.52   Mean   : 2.32   Mean   :3.27
3rd Qu.: 1.00   3rd Qu.: 1.00   3rd Qu.: 1.00   3rd Qu.: 2.00   3rd Qu.:4.00
Max.   :98.00   Max.   :99.00   Max.   :97.00   Max.   :99.00   Max.   :8.00
NA's   :395     NA's   :384     NA's   :394                     NA's   :394
     QHH2            Q41           QMARRIED        QEDUHKG
Min.   : 1.000   Min.   : 1.000   Min.   : 1.00   Min.   :1.000
1st Qu.: 2.000   1st Qu.: 2.000   1st Qu.: 1.00   1st Qu.:4.000
Median : 2.000   Median : 2.000   Median : 1.00   Median :5.000
Mean   : 3.725   Mean   : 3.359   Mean   : 3.19   Mean   :5.121
3rd Qu.: 3.000   3rd Qu.: 3.000   3rd Qu.: 4.00   3rd Qu.:7.000
Max.   :99.000   Max.   :99.000   Max.   :99.00   Max.   :8.000
                 NA's   :285                      NA's   :395
   QEDUJPN         QEDUKOR         QEDUTWN          QEDUVNM
Min.   : 3.000   Min.   : 2.000   Min.   : 3.000   Min.   : 1.00
1st Qu.: 4.000   1st Qu.: 6.000   1st Qu.: 6.000   1st Qu.: 4.00
Median : 4.000   Median : 8.000   Median : 7.000   Median : 6.00
Mean   : 8.707   Mean   : 9.087   Mean   : 6.873   Mean   : 5.47
3rd Qu.: 7.000   3rd Qu.: 8.000   3rd Qu.: 8.000   3rd Qu.: 6.00
Max.   :99.000   Max.   :99.000   Max.   :10.000   Max.   :10.00
NA's   :412      NA's   :391      NA's   :384      NA's   :394
  CHURCHEDUa      CHURCHEDUb       QPTYJPN          QPTYTWN
Min.   : 1.000   Min.   : 1.00   Min.   : 1.00   Min.   : 1.00
1st Qu.: 2.000   1st Qu.: 2.00   1st Qu.: 1.50   1st Qu.: 2.00
Median : 2.000   Median : 2.00   Median :96.00   Median : 3.00
Mean   : 1.976   Mean   : 2.35   Mean   :52.17   Mean   :44.98
```

```
 3rd Qu.: 2.000   3rd Qu.: 2.00   3rd Qu.:96.00   3rd Qu.:96.00
 Max.   :98.000   Max.   :98.00   Max.   :99.00   Max.   :99.00
                                  NA's   :412     NA's   :384
      Q50             Q51           QZIPHKG         QZIPHKG2
 Min.   :1.000   Min.   :1.000   Min.   : 1.00   Min.   : 1.000
 1st Qu.:1.000   1st Qu.:1.000   1st Qu.: 6.00   1st Qu.: 2.000
 Median :2.000   Median :1.000   Median :10.00   Median : 3.000
 Mean   :1.531   Mean   :1.083   Mean   :10.68   Mean   : 3.343
 3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:13.00   3rd Qu.: 3.000
 Max.   :2.000   Max.   :2.000   Max.   :99.00   Max.   :99.000
 NA's   :287     NA's   :410     NA's   :395     NA's   :395
    QZIPJPN2        QZIPKOR         QZIPTWN     VietnamRegion    GenderRec
 Min.   :1.000   Min.   : 1.00   Min.   : 1.0   Min.   :1.00   Min.   :  1.00
 1st Qu.:3.250   1st Qu.: 2.00   1st Qu.:11.0   1st Qu.:3.00   1st Qu.:  1.00
 Median :5.000   Median : 7.00   Median :16.0   Median :4.00   Median :  1.00
 Mean   :4.951   Mean   : 8.01   Mean   :14.7   Mean   :3.99   Mean   : 11.56
 3rd Qu.:6.750   3rd Qu.: 8.00   3rd Qu.:18.0   3rd Qu.:5.00   3rd Qu.:  2.00
 Max.   :8.000   Max.   :99.00   Max.   :99.0   Max.   :6.00   Max.   :999.00
 NA's   :412     NA's   :391     NA's   :384    NA's   :394
    Language         ISCED         UrbanVietnam
 Min.   :1.000   Min.   : 1.000   Min.   :1.00
 1st Qu.:4.000   1st Qu.: 3.000   1st Qu.:1.00
 Median :5.000   Median : 4.000   Median :2.00
 Mean   :4.877   Mean   : 4.704   Mean   :1.68
 3rd Qu.:6.000   3rd Qu.: 5.000   3rd Qu.:2.00
 Max.   :8.000   Max.   :99.000   Max.   :2.00
                                  NA's   :394
```

```r
glimpse(data_subset)
```

```
Rows: 494
Columns: 151
$ QRID          <int> 5036693, 5029082, 1514405, 4014230, 23208529, 1518511, 1~
$ weight        <dbl> 1.1983762, 0.2949189, 0.4976988, 0.7258088, 0.3814284, 0~
$ SurveyPublic  <int> 5, 5, 1, 4, 2, 1, 1, 2, 5, 6, 6, 1, 4, 6, 4, 6, 1, 1, 2,~
$ DEMONYM       <int> 5, 5, 1, 4, 2, 1, 1, 2, 5, 6, 6, 1, 4, 6, 4, 6, 1, 1, 2,~
$ Q1            <int> 2, 2, 3, 3, 2, 3, 2, 3, 98, 1, 1, 3, 3, 3, 2, 1, 3, 3, 2~
$ Q2a           <int> 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 1, 3, 3, 3, 3, 3, 3, 3, 3,~
$ Q2b           <int> 3, 3, 3, 3, 3, 3, 3, 3, 4, 3, 3, 3, 3, 3, 3, 3, 2, 3, 3,~
$ Q2c           <int> 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3,~
$ Q2d           <int> 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,~
$ Q3            <int> 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2,~
```

```
$ Q4          <int> 3, 3, 3, 3, 3, 3, 3, 3, 98, 3, 3, 3, 3, 3, 3, 3, 2, 3, 3~
$ QLEGAL      <int> 98, 2, 2, 2, 3, 2, 3, 2, 4, 2, 2, 1, 1, 2, 4, 1, 2, 2, 2~
$ Q5a         <int> 1, 3, 1, 1, 3, 2, 3, 1, 1, NA, NA, 2, 1, NA, 3, NA, 2, 3~
$ Q5b         <int> 2, 2, 1, 1, 1, 1, 2, 1, 1, NA, NA, 3, 2, NA, 2, NA, 2, 2~
$ Q5c         <int> 2, 1, 1, 2, 1, 1, 1, 1, 1, NA, NA, 2, 2, NA, 2, NA, 2, 3~
$ Q5d         <int> 1, 1, 1, 2, 1, 3, 1, 1, 1, NA, NA, 1, 2, NA, 2, NA, 2, 1~
$ QCURREL     <int> 8, 8, 8, 1, 1, 8, 1, 1, 5, 8, 1, 8, 8, 8, 8, 1, 8, 8, 1,~
$ Q6a         <int> 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 1, 2, 1, 2, 2, 1, 2, 2, 1,~
$ Q6b         <int> 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1,~
$ Q6c         <int> NA, NA, NA, NA, 2, NA, NA, 2, NA, NA, NA, NA, NA, NA, NA~
$ Q6d         <int> 2, 1, 2, 1, 2, 2, 2, 1, 1, 2, 2, 2, 1, 2, 2, 2, 2, 1, 98~
$ Q6e         <int> 2, 1, 2, 1, 2, 1, 2, 1, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 1,~
$ Q6f         <int> 2, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, 2, 1, 2, 2, 2, 2, 1, 1,~
$ QBORN       <int> NA, 2, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ Q7a         <int> 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2,~
$ Q7b         <int> 2, 1, 1, 2, 2, 2, 98, 2, 2, 2, 98, 2, 2, 2, 2, 1, 2, 1, ~
$ Q7c         <int> 2, 1, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 2, 98, 1, 1, 1, 1~
$ Q7d         <int> 2, 1, 2, 1, 2, 2, 1, 2, 1, 2, 1, 2, 1, 2, 98, 1, 1, 1, 1~
$ Q7e         <int> 2, 1, 1, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, 1, 1, 1, 1, 1, 2,~
$ Q8a         <int> NA, NA, NA, 2, 2, NA, 2, 2, NA, NA, 2, NA, NA, NA, NA, 1~
$ Q8b         <int> NA, NA, NA, 2, 2, NA, 1, 2, NA, NA, 2, NA, NA, NA, NA, 1~
$ Q8c         <int> NA, NA, NA, 2, 2, NA, 2, 2, NA, NA, 2, NA, NA, NA, NA, 1~
$ Q8d         <int> NA, NA, NA, 2, 2, NA, 1, 2, NA, NA, 2, NA, NA, NA, NA, 1~
$ Q8e         <int> NA, NA, NA, 1, 2, NA, 2, 2, NA, NA, 1, NA, NA, NA, NA, 1~
$ Q8f         <int> NA, NA, NA, 1, 2, NA, 1, 2, NA, NA, 1, NA, NA, NA, NA, 2~
$ Q8g         <int> NA, NA, NA, 1, 98, NA, 1, 2, NA, NA, 2, NA, NA, NA, NA, ~
$ Q8h         <int> NA, NA, NA, 2, 2, NA, 2, 2, NA, NA, 2, NA, NA, NA, NA, 2~
$ Q8i         <int> NA, NA, NA, 1, 1, NA, 1, 2, NA, NA, 1, NA, NA, NA, NA, 1~
$ Q8j         <int> NA, NA, NA, 2, 2, NA, 1, 2, NA, NA, 2, NA, NA, NA, NA, 9~
$ Q8k         <int> NA, NA, NA, 2, 2, NA, 1, 2, NA, NA, 2, NA, NA, NA, NA, 2~
$ Q9          <int> 3, 4, 2, 2, 3, 2, 2, 4, 1, 4, 2, 4, 2, 3, 3, 1, 4, 4, 2,~
$ Q10         <int> 1, 1, 2, 1, 98, 2, 1, 1, 1, 2, 1, 2, 2, 2, 1, 1, 2, 2, 1~
$ Q11         <int> 1, 1, 2, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 2, 1, 1, 1,~
$ Q12         <int> 2, 2, 2, 1, 2, 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 2,~
$ QCHREL      <int> 5, 10, 1, 8, 8, 5, 1, 1, 5, 8, 1, 8, 1, 8, 2, 8, 1, 7, 1~
$ Q13         <int> 1, 3, NA, NA, NA, NA, NA, NA, 3, NA, NA, NA, NA, NA, NA,~
$ Q14         <int> NA, NA, 3, NA, NA, 1, 1, NA, NA, NA, NA, 1, NA, NA, NA, ~
$ Q15a        <int> 1, 2, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2, 2, 1, 2, 2, 2, 2, 1,~
$ Q15b        <int> 1, 1, 2, 1, 2, 1, 1, 2, 1, 2, 1, 2, 1, 2, 2, 2, 2, 2, 1,~
$ Q15c        <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 2, 2,~
$ Q15d        <int> 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
$ QFIGURESa   <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 1,~
$ QFIGURESb   <int> 2, 2, 2, 2, 1, 2, 1, 1, 2, 2, 1, 2, 1, 1, 1, 1, 2, 2, 1,~
```

```
$ QFIGURESc     <int> 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 98~
$ QFIGURESd     <int> 2, 1, 2, 2, 98, 1, 2, 98, 1, 2, 1, 2, 2, 2, 2, 2, 2, 2, ~
$ QFIGURESe     <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1,~
$ QFIGURESf     <int> 2, 2, 2, 2, 1, 2, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 2, 2, 1,~
$ QFIGURESg     <int> 1, 1, 2, NA, NA, 1, 1, NA, 1, 2, 1, 2, NA, 2, NA, 2, 2, ~
$ QFIGURESh     <int> NA, NA, NA, NA, 1, NA, NA, 1, NA, NA, NA, NA, NA, NA, NA~
$ QFIGURESi     <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 1,~
$ Q16_1         <int> 1, 1, 1, 2, 1, 2, 1, 1, 2, 2, 2, 2, 2, 2, 1, 2, 1, 1, 1,~
$ Q16_2         <int> 2, 2, 2, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 1,~
$ Q16_3         <int> 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
$ Q16_98        <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
$ Q16_99        <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
$ Q17           <int> 2, 2, 2, NA, 2, NA, 2, 2, NA, NA, NA, NA, NA, NA, 2, NA,~
$ Q18           <int> 2, 4, 3, 2, 2, 2, 1, 2, 1, 1, 1, 2, 4, 1, 1, 1, 2, 3, 2,~
$ Q19           <int> 2, 2, 2, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 2, 1, 2, 2, 1,~
$ Q20a          <int> NA, NA, NA, 1, 1, 1, 1, 98, NA, 2, 2, 1, NA, 2, NA, 2, N~
$ Q20b          <int> NA, NA, NA, 1, 1, 1, 1, 1, NA, 1, 1, 1, NA, 1, NA, 1, NA~
$ Q21           <int> NA, NA, NA, 2, 2, 3, 2, 2, NA, 3, 2, 3, NA, 3, NA, 1, NA~
$ Q22a          <int> 1, 2, 1, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 2,~
$ Q22b          <int> 1, 2, 1, 1, 2, 1, 1, 1, 1, 1, 1, 2, 1, 1, 2, 1, 2, 1, 1,~
$ Q22c          <int> 1, 2, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 1,~
$ Q22d          <int> 1, 2, 1, 1, 2, 2, 1, 2, 1, 1, 1, 2, 2, 1, 2, 2, 2, 2, 2,~
$ Q23           <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2,~
$ Q24a          <int> 2, 1, 1, 2, 2, 1, 2, 2, 98, 1, 1, 2, 1, 1, 2, 1, 2, 2, 1~
$ Q24b          <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 98, 2, 1, 2, 2, 2, 2, 2, 2~
$ Q24c          <int> 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2,~
$ Q24d          <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 98, 2, 2, 1~
$ Q25a          <int> 1, 2, 2, 2, 1, 1, 2, 1, 1, 2, 1, 2, 1, 2, 1, 2, 1, 1, 1,~
$ Q25b          <int> 2, 2, 2, 2, 2, 1, 2, 2, 98, 2, 98, 2, 1, 2, 2, 2, 1, 2, ~
$ Q25c          <int> 2, 2, 2, 1, 1, 2, 2, 2, 1, 1, 1, 2, 1, 2, 2, 2, 1, 2, 2,~
$ Q26a          <int> 4, 4, 4, 4, 4, 4, 5, 5, 4, 4, 5, 3, 1, 5, 4, 4, 5, 4, 3,~
$ Q26b          <int> 2, 5, 4, 3, 4, 5, 5, 5, 4, 2, 1, 3, 1, 3, 3, 1, 5, 3, 3,~
$ Q27a          <int> 1, 2, 2, 1, 98, 2, 1, 2, 1, 2, 1, 1, 1, 2, 2, 2, 1, 2, 9~
$ Q27b          <int> 1, 2, 2, 1, 98, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 9~
$ Q27c          <int> 1, 2, 1, 1, 98, 1, 1, 98, 1, 1, 1, 1, 1, 2, 2, 1, 1, 1, ~
$ Q27d          <int> 98, 2, 1, 1, 98, 2, 1, 98, 1, 2, 98, 1, 1, 2, 2, 1, 1, 1~
$ Q27e          <int> 1, 2, 1, 2, 98, 2, 1, 2, 1, 1, 1, 2, 2, 2, 2, 1, 1, 2, 1~
$ Q27f          <int> 1, 2, 1, 1, 98, 1, 1, 1, 1, 2, 1, 1, 2, 2, 2, 2, 1, 1, 2~
$ Q27g          <int> 1, 1, 2, 2, 98, 2, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1~
$ Q27h          <int> 1, 2, 1, 1, 98, 2, 2, 1, 1, 2, 1, 1, 1, 2, 2, 2, 1, 2, 1~
$ Q27i          <int> 1, 2, 1, 1, 98, 2, 1, 1, 1, 2, 1, 1, 2, 2, 2, 1, 1, 1, 1~
$ Q28           <int> 6, 6, 6, 7, 1, 7, 5, 7, 98, 7, 5, 7, 5, 7, 7, 7, 7, 7, 6~
$ Q29           <int> 2, 2, 2, 2, 1, 2, 1, 2, 1, 1, 1, 2, 2, 1, 2, 1, 2, 2, 1,~
```

```
$ Q30           <int> 2, 1, 2, 2, 2, 2, 2, 2, 2, 3, 2, 2, 2, 2, 2, 2, 1, 1, 2,~
$ Q31           <int> 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 1, 2, 2, 2, 1, 2, 1,~
$ Q33           <int> 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2,~
$ Q35           <int> 1, 1, 1, 2, 2, 1, 2, 98, 1, 3, 2, 2, 2, 2, 1, 1, 2, 2, 1~
$ Q36           <int> 1, 1, 2, 1, 1, 2, 1, 98, 2, 2, 2, 2, 1, 1, 1, 2, 1, 2, 1~
$ Q37           <int> 1, 1, 1, 2, 2, 1, 1, 2, 3, NA, NA, 1, 1, NA, 1, NA, 1, 1~
$ Q38a          <int> 2, 2, 2, 1, 1, 1, 1, 2, 2, 1, 1, 1, 2, 2, 2, 1, 2, 2, 2,~
$ Q38b          <int> 98, 2, 2, 1, 2, 1, 1, 2, 2, 2, 98, 2, 2, 2, 2, 2, 2, 2, ~
$ Q38c          <int> 2, 2, 2, 1, 2, 1, 1, 2, 2, NA, NA, 1, 1, NA, 2, NA, 2, 2~
$ Q39a          <int> 2, 2, 3, 1, 3, 3, 3, 4, 2, 1, 2, 4, 1, 2, 1, 1, 2, 2, 2,~
$ Q39b          <int> 3, 2, 2, 2, 2, 2, 3, 2, 4, 4, 2, 2, 4, 4, 3, 4, 4, 3, 2,~
$ Q39c          <int> 2, 2, 2, 2, 3, 2, 2, 3, 1, 1, 1, 4, 1, 2, 1, 1, 1, 1, 2,~
$ Q39d          <int> 3, 3, 2, 2, 2, 4, 4, 4, 4, 4, 3, 2, 3, 4, 4, 4, 4, 4, 3,~
$ QABORT        <int> 2, 3, 2, 1, 2, 3, 2, 98, 98, 3, 2, 4, 1, 4, 2, 2, 4, 2, ~
$ Q40a          <int> 1, 2, 2, 1, 2, 2, 1, 4, 4, NA, NA, 2, 1, NA, 1, NA, 2, 2~
$ Q40b          <int> 4, 4, 4, 4, 4, 4, 4, 3, 4, NA, NA, 4, 4, NA, 4, NA, 4, 4~
$ Q40c          <int> 4, 4, 2, 2, 2, 4, 4, 4, 4, NA, NA, 4, 4, NA, 4, NA, 4, 4~
$ Q40d          <int> 4, 2, 2, 1, 1, 2, 4, 2, 4, NA, NA, 4, 2, NA, 4, NA, 2, 4~
$ Q40e          <int> 4, 4, 3, 4, 4, 4, 4, 2, 98, NA, NA, 4, 4, NA, 4, NA, 4, ~
$ QATTEND       <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QAGE          <int> 20, 47, 21, 65, 62, 46, 72, 44, 60, 23, 24, 24, 47, 40, ~
$ QAGE_NRHKG    <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QAGE_NRJPN    <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QAGE_NRKOR    <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QAGE_NRTWNrec <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QAGE_NRVNM    <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
$ QETHHKG       <int> NA, NA, 1, NA, NA, 1, 1, NA, NA, NA, NA, 1, NA, NA, NA, ~
$ QETHTWN       <int> 2, 1, NA, NA, NA, NA, NA, NA, 1, NA, NA, NA, NA, NA, NA,~
$ QETHVNM       <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, 1, 1, NA, NA, 1, NA,~
$ QFERT         <int> 1, 0, 0, 3, 0, 0, 2, 3, 3, 0, 2, 0, 0, 2, 0, 2, 0, 0, 3,~
$ QHH1          <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, 3, 4, NA, NA, 5, NA,~
$ QHH2          <int> 4, 1, 4, 3, 1, 5, 3, 3, 1, 2, 2, 3, 2, 4, 2, 3, 2, 1, 5,~
$ Q41           <int> 3, 3, 2, NA, NA, 2, 3, NA, 4, NA, NA, 2, NA, NA, NA, NA,~
$ QMARRIED      <int> 1, 4, 4, 1, 2, 4, 5, 1, 3, 4, 1, 4, 4, 1, 1, 1, 2, 4, 1,~
$ QEDUHKG       <int> NA, NA, 5, NA, NA, 6, 2, NA, NA, NA, NA, 7, NA, NA, NA, ~
$ QEDUJPN       <int> NA, NA, NA, NA, 4, NA, NA, 4, NA, NA, NA, NA, NA, NA, NA~
$ QEDUKOR       <int> NA, NA, NA, 6, NA, NA, NA, NA, NA, NA, NA, NA, 9, NA, 8,~
$ QEDUTWN       <int> 6, 8, NA, NA, NA, NA, NA, NA, 6, NA, NA, NA, NA, NA, NA,~
$ QEDUVNM       <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, 6, 4, NA, NA, 6, NA,~
$ CHURCHEDUa    <int> 2, 2, 1, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 1, 1,~
$ CHURCHEDUb    <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2,~
$ QPTYJPN       <int> NA, NA, NA, NA, 96, NA, NA, 98, NA, NA, NA, NA, NA, NA, ~
$ QPTYTWN       <int> 96, 2, NA, NA, NA, NA, NA, NA, 2, NA, NA, NA, NA, NA, NA~
```

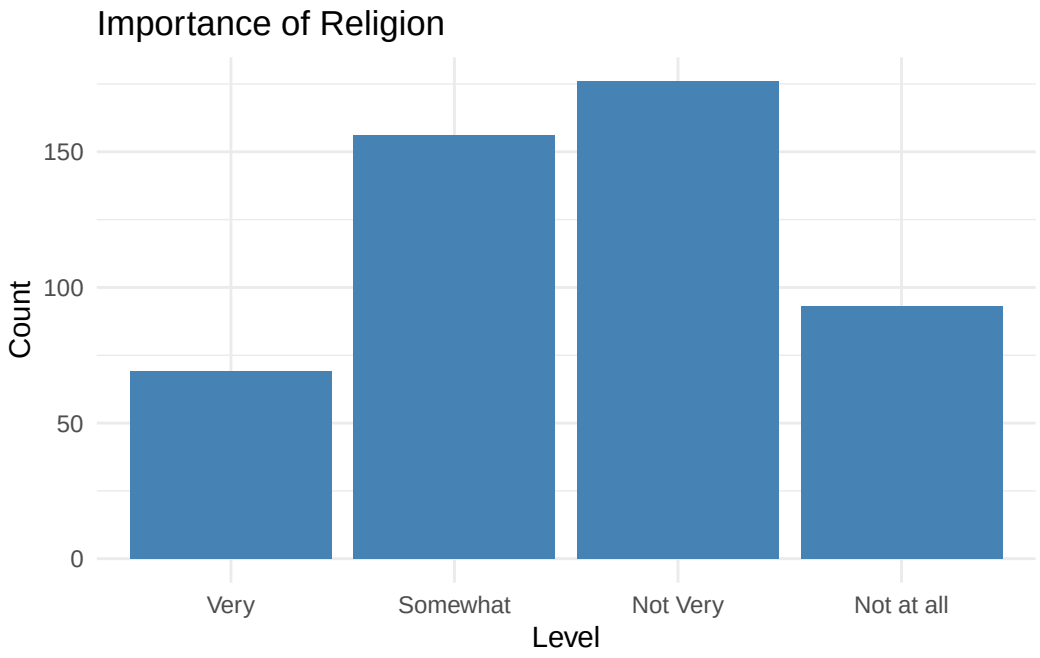
```
$ Q50           <int> 2, NA, 2, NA, 1, 2, NA, 2, NA, NA, NA, 2, NA, NA, NA, NA~
$ Q51           <int> NA, 1, NA, NA, NA, NA, 1, NA, 1, NA, NA, NA, NA, NA, NA,~
$ QZIPHKG       <int> NA, NA, 10, NA, NA, 7, 15, NA, NA, NA, NA, 14, NA, NA, N~
$ QZIPHKG2      <int> NA, NA, 3, NA, NA, 2, 3, NA, NA, NA, NA, 3, NA, NA, NA, ~
$ QZIPJPN2      <int> NA, NA, NA, NA, 8, NA, NA, 4, NA, NA, NA, NA, NA, NA, NA~
$ QZIPKOR       <int> NA, NA, NA, 7, NA, NA, NA, NA, NA, NA, NA, NA, 3, NA, 7,~
$ QZIPTWN       <int> 16, 7, NA, NA, NA, NA, NA, NA, 22, NA, NA, NA, NA, NA, N~
$ VietnamRegion <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, 5, 5, NA, NA, 4, NA,~
$ GenderRec     <int> 2, 2, 2, 1, 1, 1, 2, 1, 1, 2, 2, 2, 1, 1, 1, 1, 2, 2, 1,~
$ Language      <int> 6, 6, 1, 5, 4, 1, 1, 4, 6, 8, 8, 1, 5, 8, 5, 8, 1, 1, 4,~
$ ISCED         <int> 3, 5, 4, 3, 3, 5, 1, 3, 3, 3, 2, 5, 5, 3, 5, 3, 3, 2, 3,~
$ UrbanVietnam  <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, 2, 2, NA, NA, 2, NA,~
```

**Exploratory Plots:**

```
library(tidyverse)
library(scales)
```

```
Attaching package: 'scales'

The following object is masked from 'package:purrr':

    discard

The following object is masked from 'package:readr':

    col_factor
```

```
# Bar Plot: Importance of Religion
ggplot(data_subset, aes(x = factor(Q9,
                                   levels = c(1, 2, 3, 4, 98, 99),
                                   labels = c("Very", "Somewhat", "Not Very", "Not at all", "
  geom_bar(fill = "steelblue") +
  labs(title = "Importance of Religion", x = "Level", y = "Count") +
  theme_minimal()
```

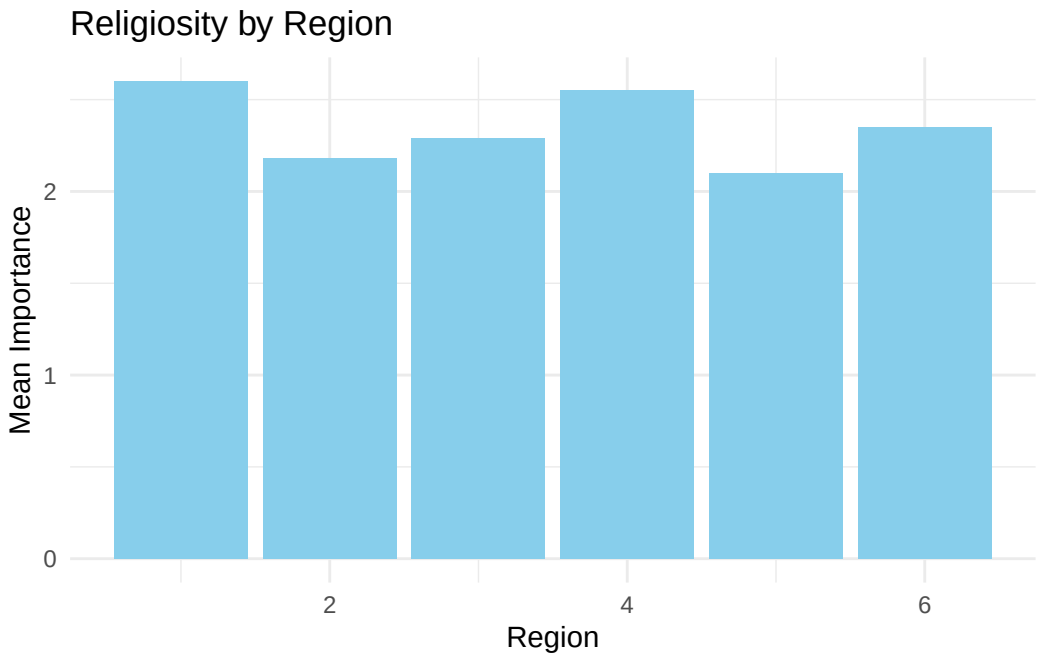



```
# Correlation Between Age and Importance of Religion
correlation <- cor(data_subset$QAGE, as.numeric(data_subset$Q9), use = "complete.obs")
print(paste("Correlation between Age and Religion Importance:", correlation))
```

```
[1] "Correlation between Age and Religion Importance: 0.0480425255982501"
```

```
# Bar Plot: Religiosity by Region
data_subset %>%
  group_by(VietnamRegion) %>%
  summarize(mean_religion_importance = mean(as.numeric(Q9), na.rm = TRUE)) %>%
  ggplot(aes(x = VietnamRegion, y = mean_religion_importance)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "Religiosity by Region", x = "Region", y = "Mean Importance") +
  theme_minimal()
```

```
Warning: Removed 1 row containing missing values or values outside the scale range
(`geom_bar()`).
```




```
# Chi-Square Test: Religion Importance and Region
chisq_result <- chisq.test(table(data_subset$Q9, data_subset$VietnamRegion))
```

```
Warning in chisq.test(table(data_subset$Q9, data_subset$VietnamRegion)):
Chi-squared approximation may be incorrect
```

```
print("Chi-Square Test Results:")
```

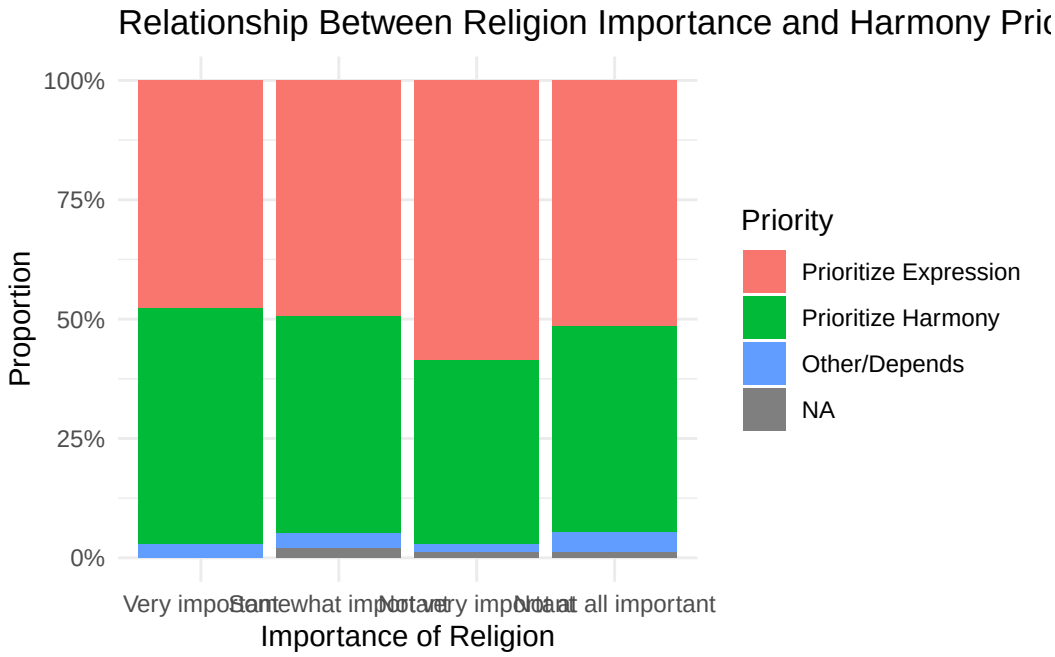
```
[1] "Chi-Square Test Results:"
```

```
print(chisq_result)
```

```

    Pearson's Chi-squared test

data:  table(data_subset$Q9, data_subset$VietnamRegion)
X-squared = 12.563, df = 15, p-value = 0.636
```

```
# Relationship Between Religion Importance and Harmony Priority
data_subset <- data_subset %>%
  mutate(
```

```
    religion_importance = factor(Q9,
                                 levels = c(1, 2, 3, 4, 98, 99),
                                 labels = c("Very important", "Somewhat important", "Not very
    harmony_priority = factor(Q36,
                              levels = c(1, 2, 3),
                              labels = c("Prioritize Expression", "Prioritize Harmony", "Othe
  )

ggplot(data_subset, aes(x = religion_importance, fill = harmony_priority)) +
  geom_bar(position = "fill") +
  scale_y_continuous(labels = percent_format()) +
  labs(
    title = "Relationship Between Religion Importance and Harmony Priority",
    x = "Importance of Religion",
    y = "Proportion",
    fill = "Priority"
  ) +
  theme_minimal()
```



## Team Charter

**When will you meet as a team to work on the project components? Will these meetings be held in person or virtually?**

We will do both as needed, on an ad-hoc basis.

**What is your group policy on missing team meetings (e.g., how much advance notice should be provided)?**

One day in advance. The person missing the meeting will get a slap on the wrist and do more work. :D

**How will your team communicate (email, Slack, text messages)? What is your policy on appropriate response time (within a certain number of hours? Nights/weekends?)?**

We will communicate via whatsapp, responses should be in the same day.