Analysis of Kenyan Social Values from World Values Survey

Installing packages

```
#install.packages("tidyverse")
#install.packages("gtsummary")
#install.packages("gt")
#install.packages("here")
#install.packages("usethis")
```

Libraries

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.1     v stringr   1.6.0
v ggplot2   4.0.0     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.2.0
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to beco
```

```
library(gtsummary)
library(gt)
library(here)
```

```
here() starts at /Users/johnmusalia/Documents/git-demo
```

```
library(usethis)
```

Reading Data

```
wvs <- read_csv(here("Data","WVS.csv"))
```

```
Rows: 94278 Columns: 606
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (12): version, doi, B_COUNTRY_ALPHA, C_COW_ALPHA, LNGE_ISO, X002_02B, V...
dbl (594): A_WAVE, A_YEAR, A_STUDY, B_COUNTRY, C_COW_NUM, D_INTERVIEW, S007,...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
wvs <- rename_with(wvs, tolower)
#wvs <- wvs |>
#    mutate(across(everything(), ~replace(., .x < 0, NA)))
wvs |> count(q1)
```

```
# A tibble: 7 x 2
     q1     n
  <dbl> <int>
1    -5    67
2    -2    61
3    -1    27
4     1 84694
5     2  8315
6     3   896
7     4   218
```

Checking what kind of variables to work on

```r
wvs <- wvs |>
    filter(b_country_alpha == "KEN") |>
    mutate(sex = ifelse(q260 == 1, "Male", "Female"),
           Friends = case_when(q2 ==1 ~ "Very Important",
                               q2 == 2 ~ "Rather Important",
                               q2 == 3 ~ " Not Very Important",
                               q2 == 4 ~ "Not at all Important"),
           friends = factor(q2,
                            levels = c(1,2,3,4),
                            labels = c("Very Important",
                                       "Rather Important",
                                       "Not Very Important",
                                       "Not at all Important")),
           leisure = factor(q3,
                            levels = c(1,2,3,4),
                            labels = c("Very Important",
                                       "Rather Important",
                                       "Not Very Important",
                                       "Not at all Important")),
           politics = factor(q4,
                             levels = c(1,2,3,4),
                             labels = c("Very Important",
                                        "Rather Important",
                                        "Not Very Important",
                                        "Not at all Important")),
           work = factor(q5,
                         levels = c(1,2,3,4),
                         labels = c("Very Important",
                                    "Rather Important",
                                    "Not Very Important",
                                    "Not at all Important")),
           religion = factor(q6,
                             levels = c(1,2,3,4),
                             labels = c("Very Important",
                                        "Rather Important",
                                        "Not Very Important",
                                        "Not at all Important")),
           residence = factor(h_urbrural,
                              levels = c(1,2),
                              labels = c("Urban", "Rural")),
          happy = factor(q46,
                         levels = c(1,2,3,4),
```

```
                          labels = c("Very Happy",
                                     "Quite Happy",
                                     "Not Very Happy",
                                     "Not at all Happy"))
            )
wvs |> count(Friends)
```

```
# A tibble: 5 x 2
  Friends                    n
  <chr>                  <int>
1 " Not Very Important"    109
2 "Not at all Important"    16
3 "Rather Important"       445
4 "Very Important"         694
5  <NA>                      2
```

```
wvs |> count(leisure)
```

```
# A tibble: 5 x 2
  leisure                  n
  <fct>                <int>
1 Very Important         602
2 Rather Important       393
3 Not Very Important     221
4 Not at all Important    46
5 <NA>                     4
```

```
wvs |> count(politics)
```

```
# A tibble: 5 x 2
  politics                 n
  <fct>                <int>
1 Very Important         296
2 Rather Important       276
3 Not Very Important     377
4 Not at all Important   298
5 <NA>                    19
```

```
wvs |> count(work)
```

```
# A tibble: 5 x 2
  work                    n
  <fct>               <int>
1 Very Important       1128
2 Rather Important      103
3 Not Very Important     21
4 Not at all Important   13
5 <NA>                    1
```

```
wvs |> count(religion)
```

```
# A tibble: 5 x 2
  religion                n
  <fct>               <int>
1 Very Important       1097
2 Rather Important      110
3 Not Very Important     39
4 Not at all Important   19
5 <NA>                    1
```

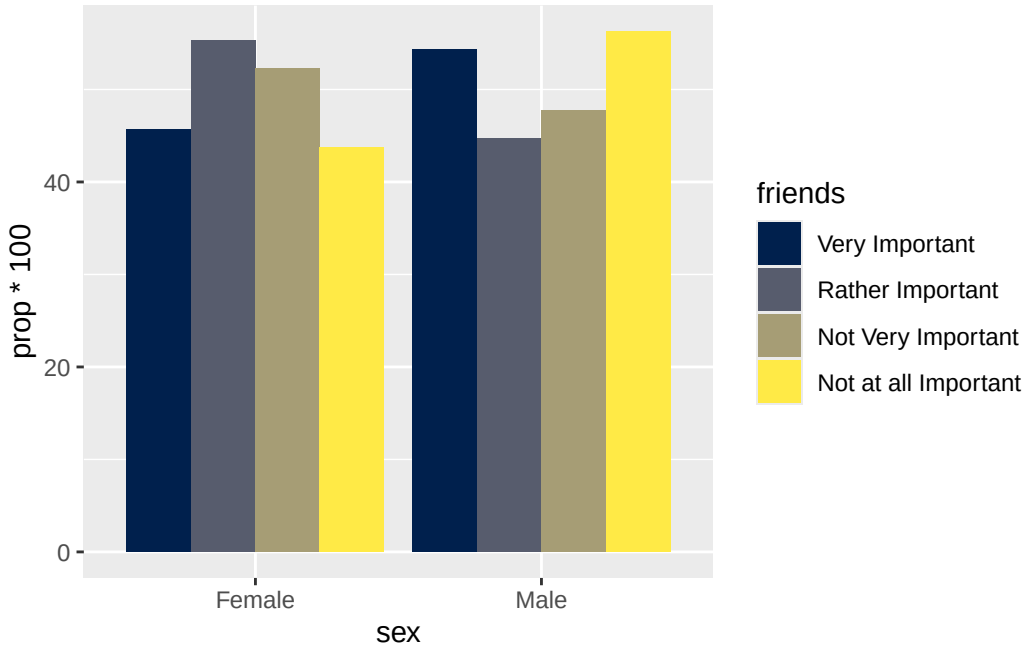
```
wvs |> count(residence)
```

```
# A tibble: 2 x 2
  residence     n
  <fct>     <int>
1 Urban       719
2 Rural       547
```

```
wvs |> count(happy)
```

```
# A tibble: 5 x 2
  happy               n
  <fct>           <int>
1 Very Happy        634
2 Quite Happy       415
3 Not Very Happy    170
4 Not at all Happy   34
5 <NA>               13
```

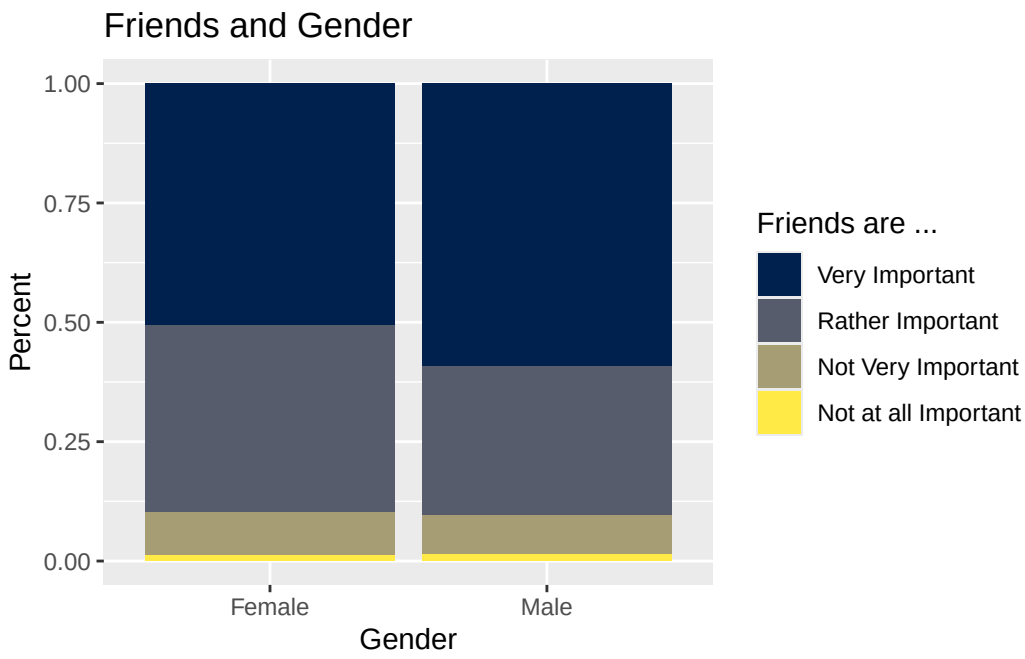
```
#Friends
wvs |>
    drop_na(sex, Friends) |>
    ggplot(aes(x=sex,
               y =after_stat(prop*100), group= friends, fill=friends)) +
    geom_bar(position = "dodge") +
    scale_fill_viridis_d(option="E")
```




```
#wvs |> count(q260)
```

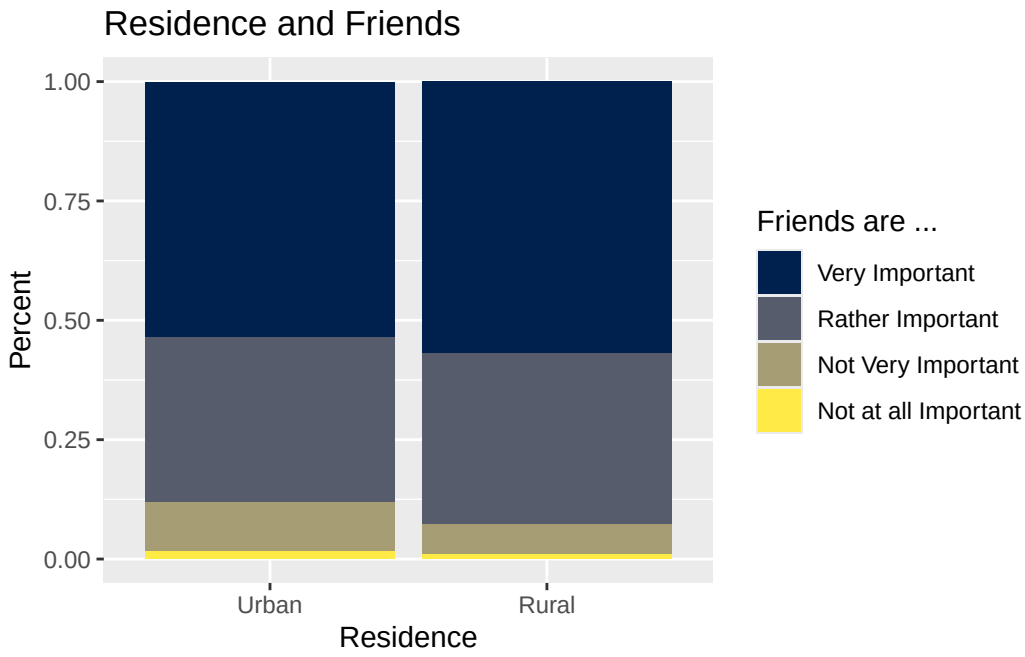
Friends

```
wvs |>
    drop_na(sex, friends) |>
    ggplot(aes(x = sex, fill = friends)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Friends and Gender",
y = "Percent",
x = "Gender",
fill = "Friends are ...")
```



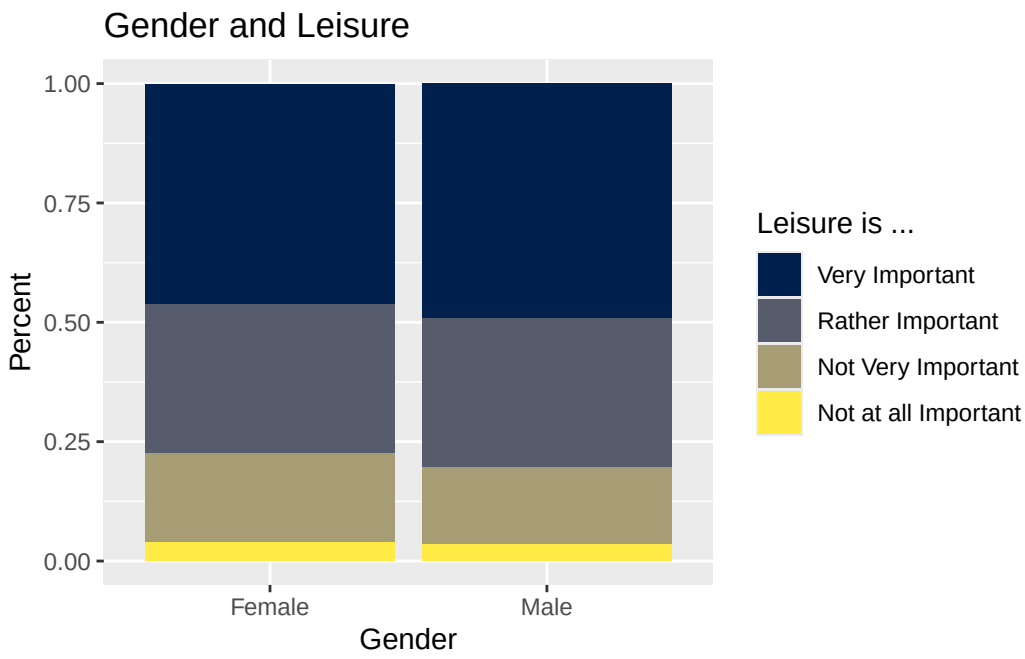
Friends by Residence

```
wvs |>
    drop_na(residence, friends) |>
    ggplot(aes(x = residence, fill = friends)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Residence and Friends",
y = "Percent",
x = "Residence",
fill = "Friends are ...")
```



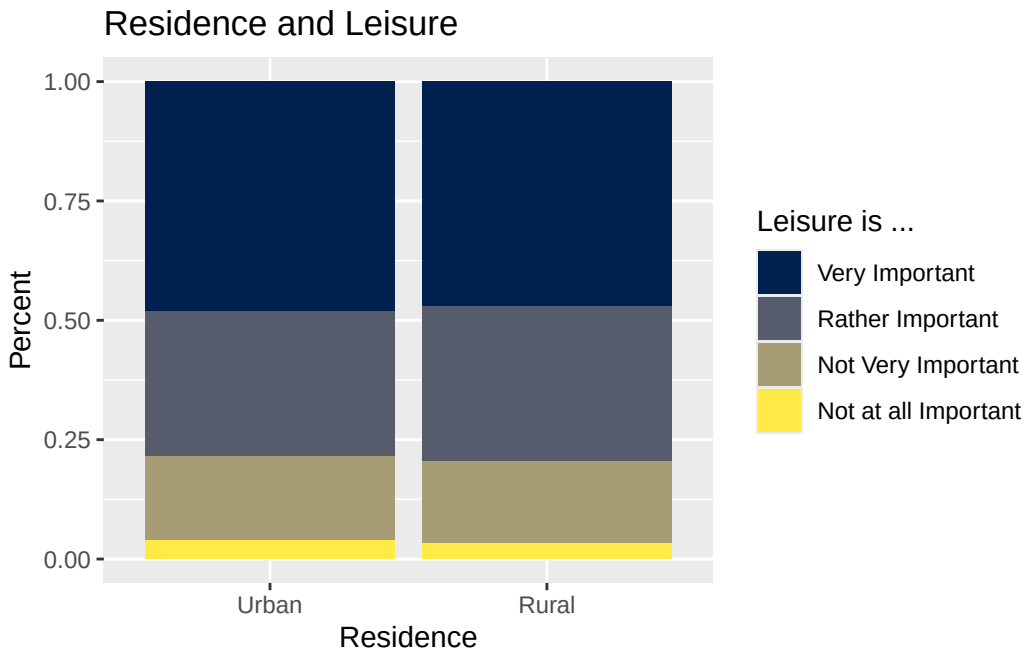
Leisure

```
wvs |>
    drop_na(sex, leisure) |>
    ggplot(aes(x = sex, fill = leisure)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Gender and Leisure",
y = "Percent",
x = "Gender",
fill = "Leisure is ...")
```



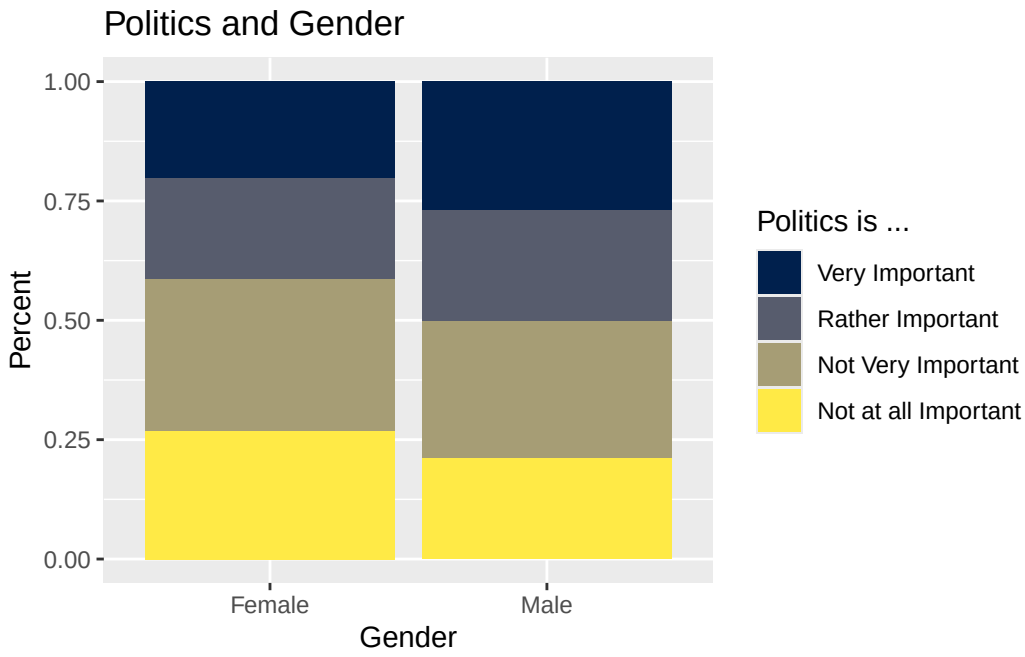
Leisure with Residence

```
wvs |>
    drop_na(residence, leisure) |>
    ggplot(aes(x = residence, fill = leisure)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Residence and Leisure",
y = "Percent",
x = "Residence",
fill = "Leisure is ...")
```



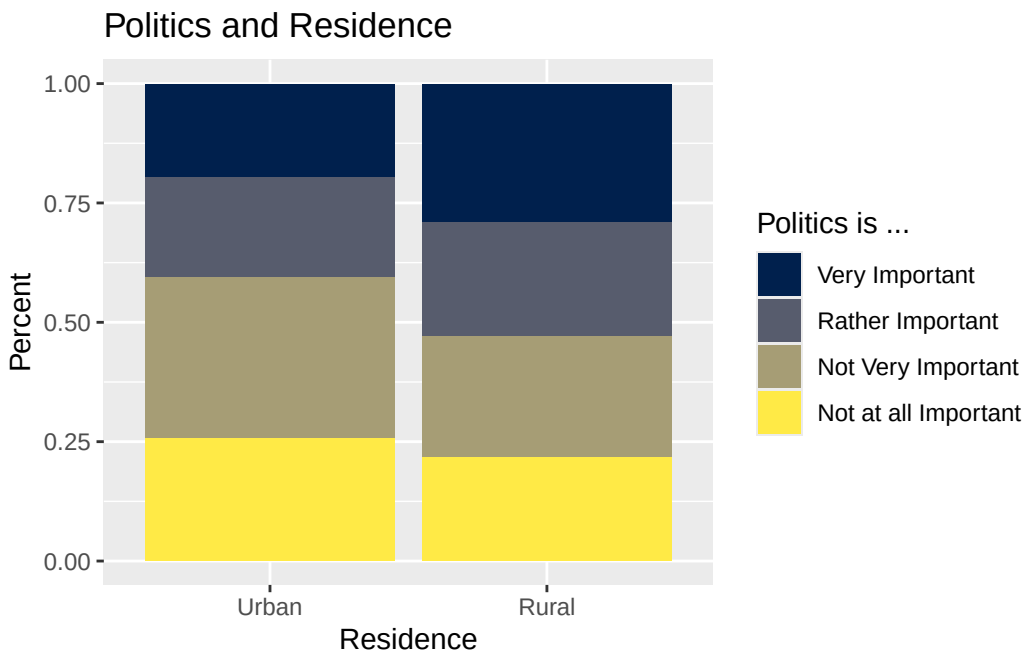
Politics

```
wvs |>
    drop_na(sex, politics) |>
    ggplot(aes(x = sex, fill = politics)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Politics and Gender",
y = "Percent",
x = "Gender",
fill = "Politics is ...")
```



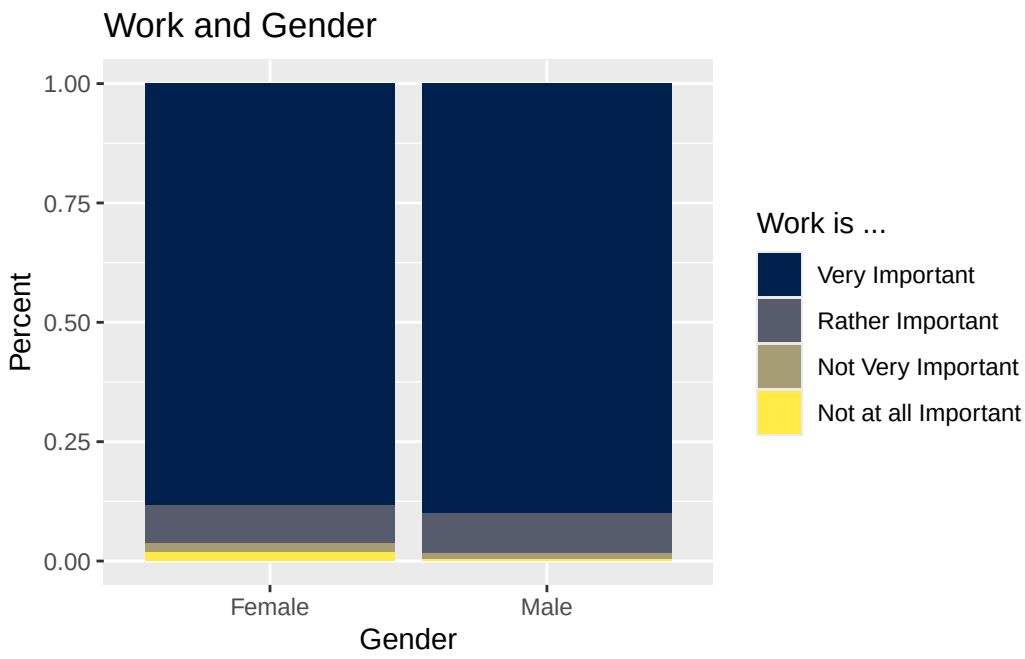
Politics and Residence

```
wvs |>
    drop_na(residence, politics) |>
    ggplot(aes(x = residence, fill = politics)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Politics and Residence",
y = "Percent",
x = "Residence",
fill = "Politics is ...")
```



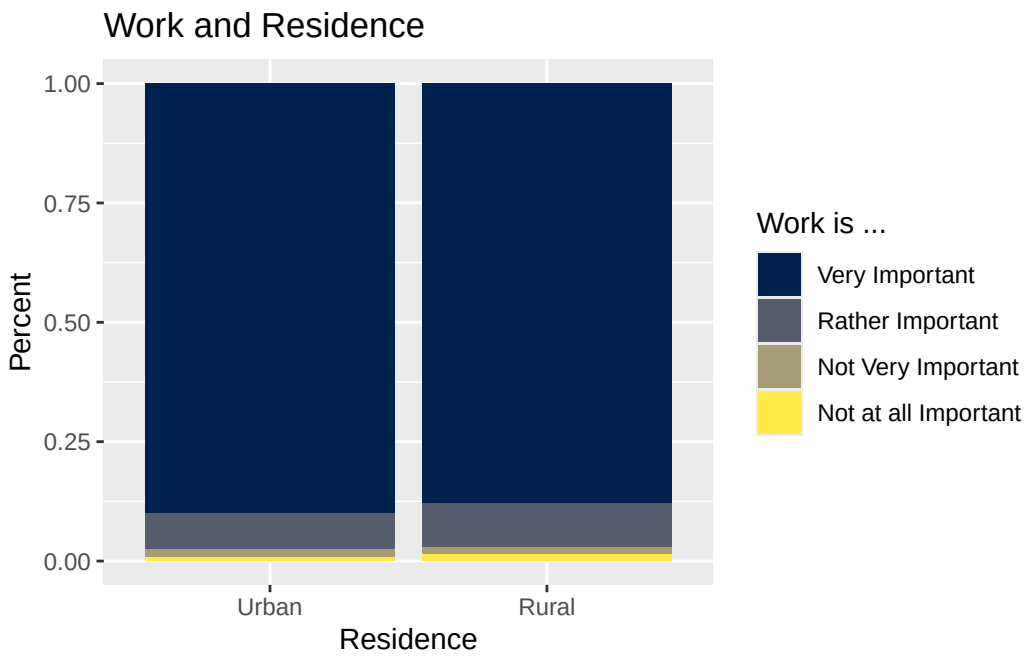
Work

```
wvs |>
    drop_na(sex, work) |>
    ggplot(aes(x = sex, fill = work)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Work and Gender",
y = "Percent",
x = "Gender",
fill = "Work is ...")
```



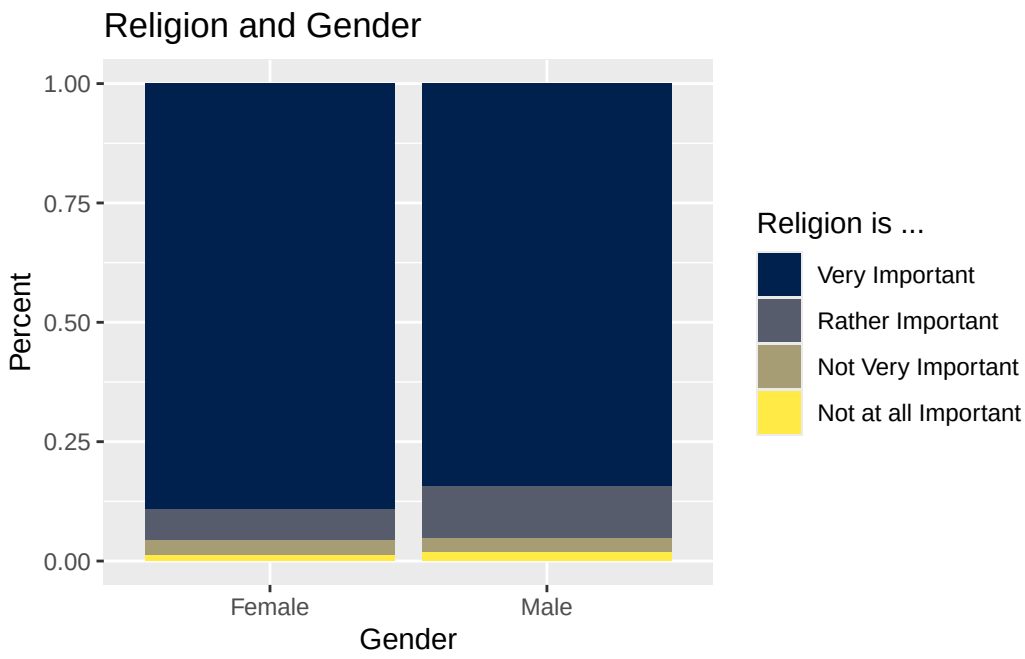
Work and Residence

```
wvs |>
    drop_na(residence, work) |>
    ggplot(aes(x = residence, fill = work)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Work and Residence",
y = "Percent",
x = "Residence",
fill = "Work is ...")
```



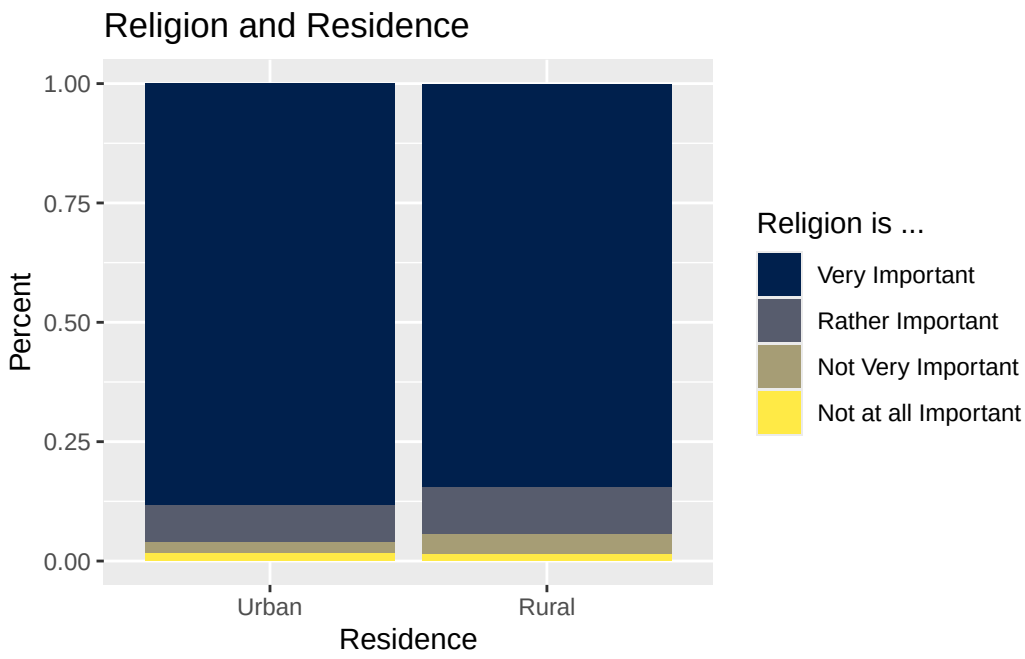
Religion

```
wvs |>
    drop_na(sex, religion) |>
    ggplot(aes(x = sex, fill = religion)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Religion and Gender",
y = "Percent",
x = "Gender",
fill = "Religion is ...")
```



Religion and Residence

```
wvs |>
    drop_na(residence, religion) |>
    ggplot(aes(x = residence, fill = religion)) +
    geom_bar(position = "fill") +
scale_fill_viridis_d(option = "E") +
labs(title = "Religion and Residence",
y = "Percent",
x = "Residence",
fill = "Religion is ...")
```


Religion and Residence
1.00
0.75
0.50
0.25
0.00
Percent
Urban
Rural
Residence
Religion is ...
Very Important
Rather Important
Not Very Important
Not at all Important