

Final Report

due November 16, 2021 by 11:59 PM

Ashley Bae, Gi Chun, Jay Nolt

November 2021

Load Packages

```
library(tidyverse)
library(readxl)
library(dbplyr)
library(ggplot2)
```

Load Data

```
load(file = "~/sta198/Jackie-Fan-Club/data/ICPSR_34363/DS0001/34363-0001-Data.rda")
#rename file
data <- da34363.0001
```

Data Citation: Battle, Juan, Pastrana, Antonio Jay, and Daniels, Jessie. Social Justice Sexuality Project: 2010 National Survey, including Puerto Rico. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-08-09. <https://doi.org/10.3886/ICPSR34363.v1>

```
data_filtered <- data %>%
  select(Q15B,
         Q15C,
         Q17A,
         Q17B,
         Q18A1:Q18A5,
         Q18C,
         Q18G,
         Q19A1:Q19A7,
         Q22A,
         Q22B,
         Q25)

# remove descriptions and other attributes
data_filtered2 <- lapply(data, function(x) {attributes(x) <- NULL; x}) %>%
  as.data.frame() %>%
  select(Q15B,
         Q15C,
         Q17A,
         Q17B,
         Q18A1:Q18A5,
         Q18C,
```

```
Q18G,  
Q19A1:Q19A7,  
Q22A,  
Q22B,  
Q25)
```

```
data <- data_filtered %>%  
  select(Q17A,  
    Q17B,  
    Q18A1:Q18A5,  
    Q18C,  
    Q18G,  
    Q19A1:Q19A7,  
    Q22A,  
    Q22B,  
    Q25) %>%  
  rename(healthInsureAcc = Q17A) %>%  
  rename(healthProvideAcc = Q17B) %>%  
  rename(male = Q18A1) %>%  
  rename(female = Q18A2) %>%  
  rename(m2f = Q18A3) %>%  
  rename(f2m = Q18A4) %>%  
  rename(genderOther = Q18A5) %>%  
  rename(sexuality = Q18C) %>%  
  mutate(age = 2021 - Q18G) %>%  
  rename(black = Q19A1) %>%  
  rename(hispanic = Q19A2) %>%  
  rename(asian = Q19A3) %>%  
  rename(native = Q19A4) %>%  
  rename(white = Q19A5) %>%  
  rename(multi = Q19A6) %>%  
  rename(raceOther = Q19A7) %>%  
  rename(edu = Q22A) %>%  
  rename(income = Q22B) %>%  
  rename(assessHealth = Q25)
```

```
data <- data %>%  
  mutate(male = case_when(  
    male == "(1) Yes" ~ "male",  
    is.na(male) ~ "",  
    TRUE ~ ""),  
    female = case_when(  
    female == "(1) Yes" ~ "female",  
    is.na(female) ~ "",  
    TRUE ~ ""),  
    m2f = case_when(  
    m2f == "(1) Yes" ~ "m2f",  
    is.na(m2f) ~ "",  
    TRUE ~ ""),  
    f2m = case_when(  
    f2m == "(1) Yes" ~ "f2m",  
    is.na(f2m) ~ "",  
    TRUE ~ ""),  
    genderOther = case_when(  
    genderOther == "(1) Yes" ~ "genderOther",  
    is.na(genderOther) ~ "",  
    TRUE ~ ""))
```

```

genderOther == "(1) Yes" ~ "Other",
is.na(genderOther) ~ "",
TRUE ~ ""),
gender = ""
)

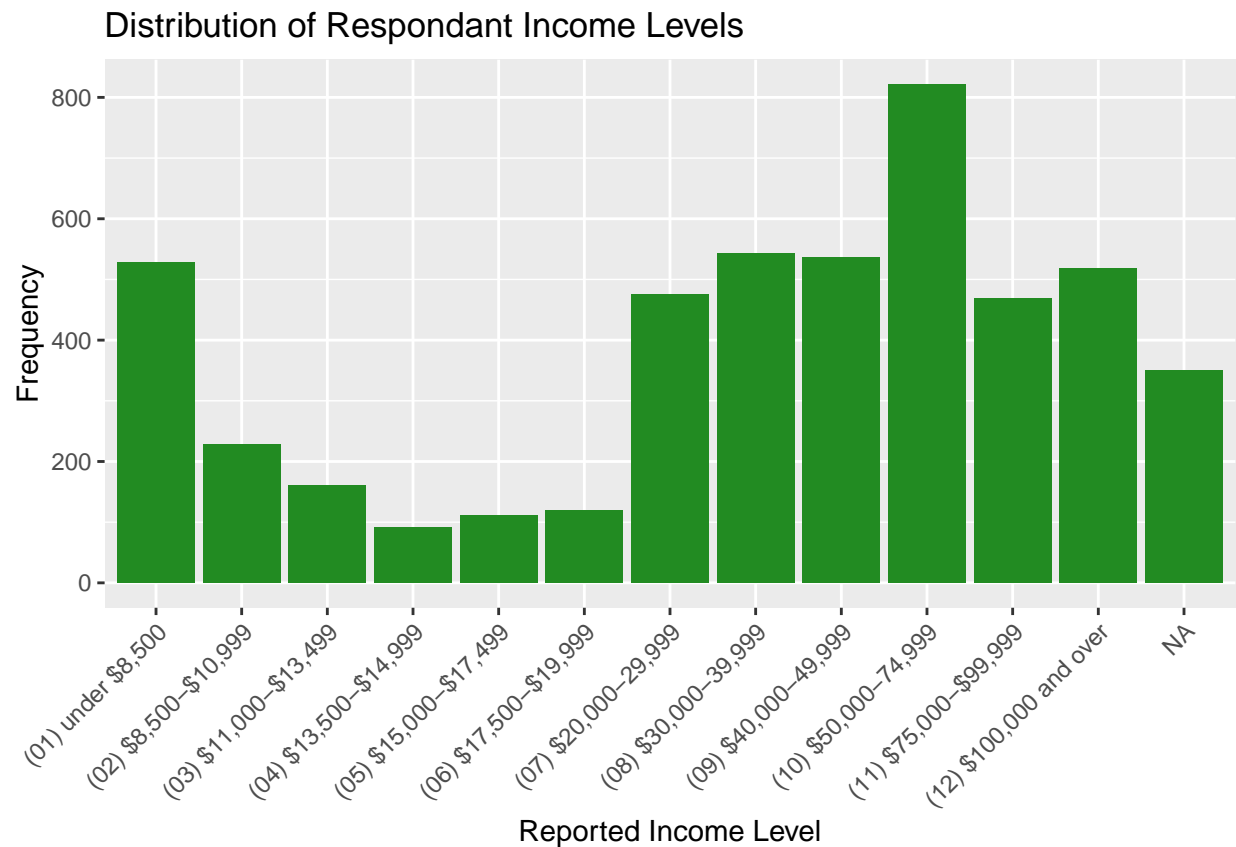
data$gender <- paste(data$male, data$female)
data$gender <- paste(data$gender, data$m2f)
data$gender <- paste(data$gender, data$f2m)
data$gender <- paste(data$gender, data$genderOther)

data$gender <- trimws(data$gender)

data <- data %>%
mutate(gender = case_when(
  gender == "male" ~ "Male",
  gender == "female" ~ "Female",
  gender == "m2f" ~ "Transgender: Male to Female",
  gender == "f2m" ~ "Transgender: Female to Male",
  TRUE ~ "Other")
)

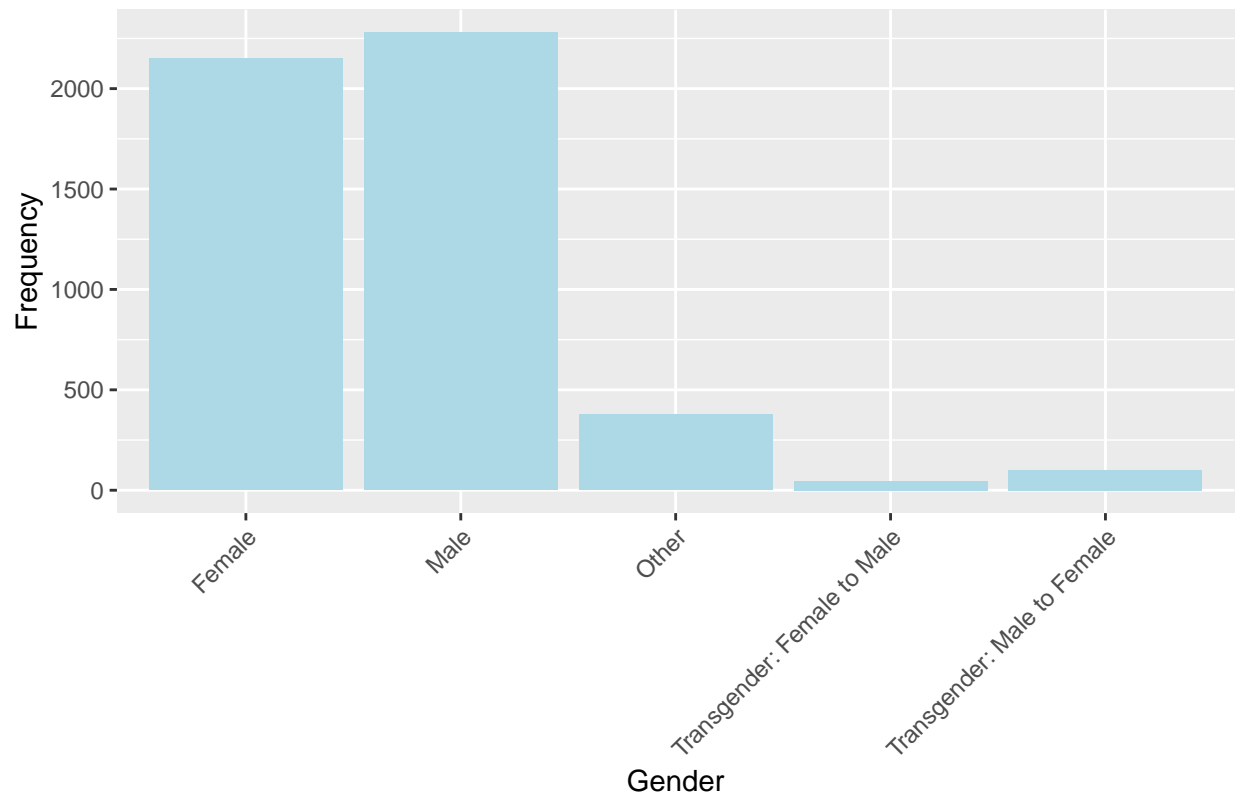
ggplot(data, aes(x = income)) +
  geom_bar(fill = "#228b22") +
  labs (x = "Reported Income Level",
        y = "Frequency",
        title = "Distribution of Respondant Income Levels") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))

```



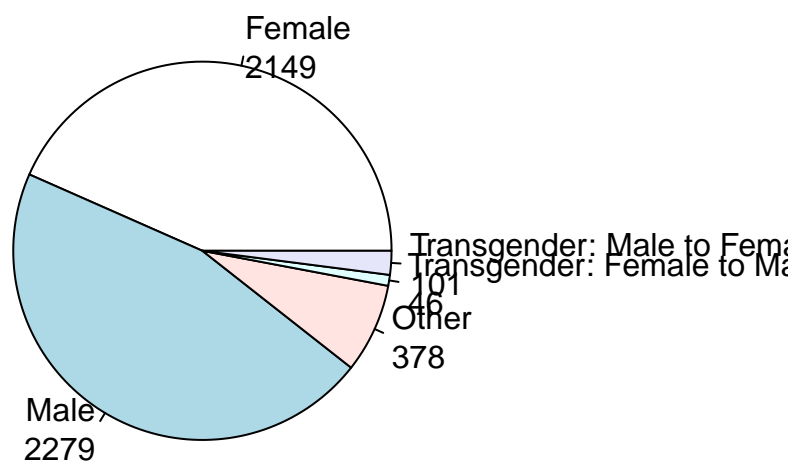
```
ggplot(data, aes(x = gender)) +
  geom_bar(fill = "light blue") +
  labs (x = "Gender",
        y = "Frequency",
        title = "Distribution of Respondant Gender") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```

Distribution of Respondant Gender

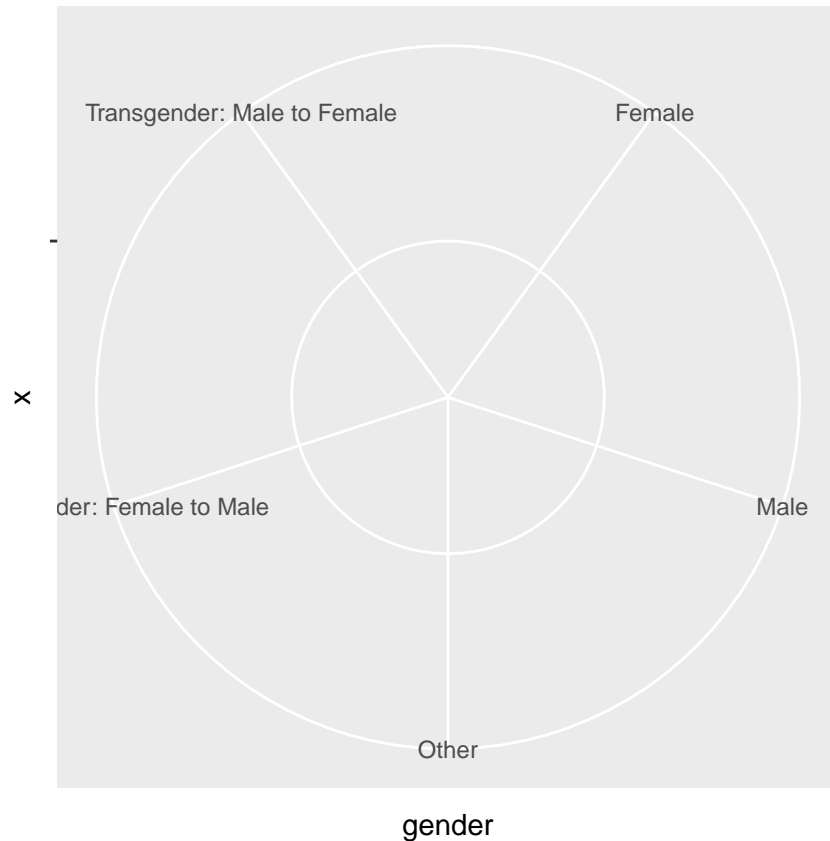


```
# Pie Chart with Percentages
gendertable <- table(data$gender)
lbls <- paste(names(gendertable), "\n", gendertable, sep="")
pie(gendertable, labels = lbls,
    main="Pie Chart of Gender Distribution of Respondents")
```

Pie Chart of Gender Distribution of Respondents



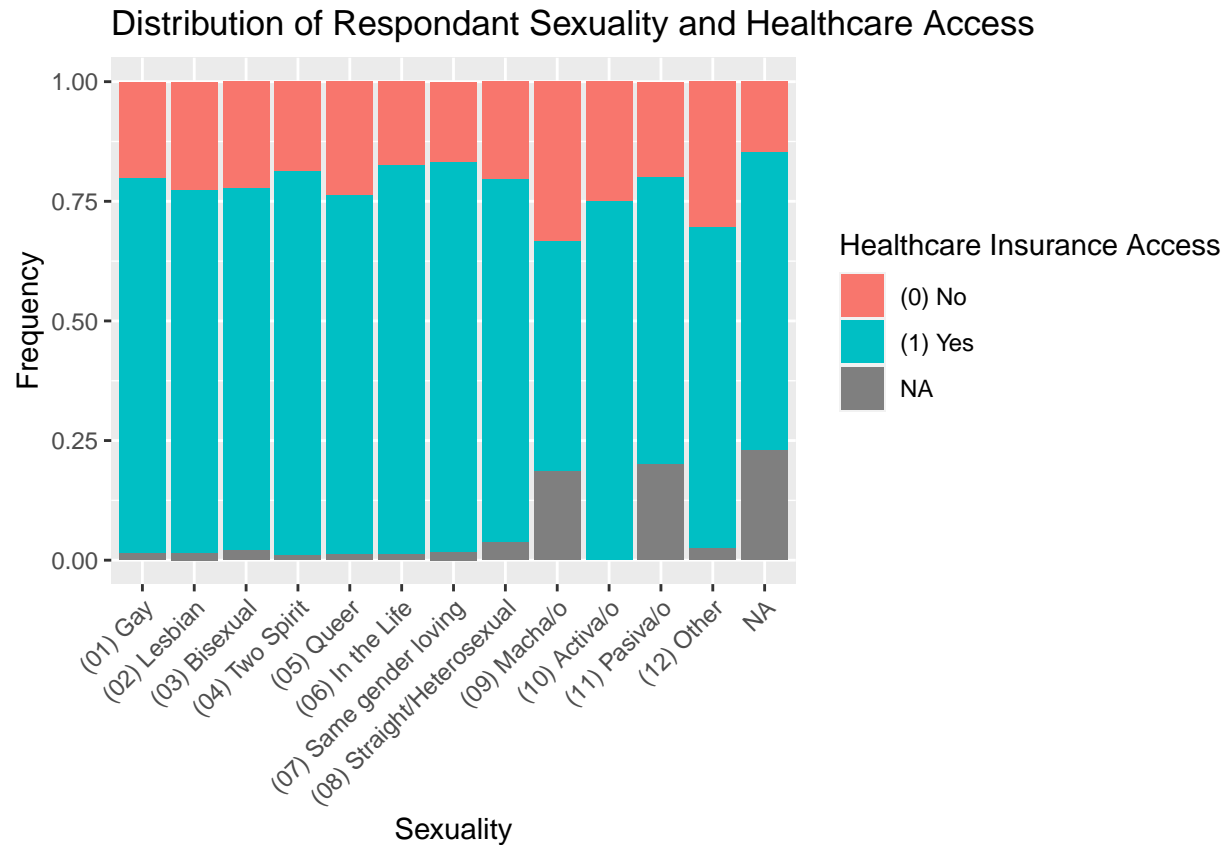
```
# attempt at ggplot2 way...  
ggplot(data, aes(x = "", y = gender, fill = "gender")) +  
  coord_polar(theta = "y")
```



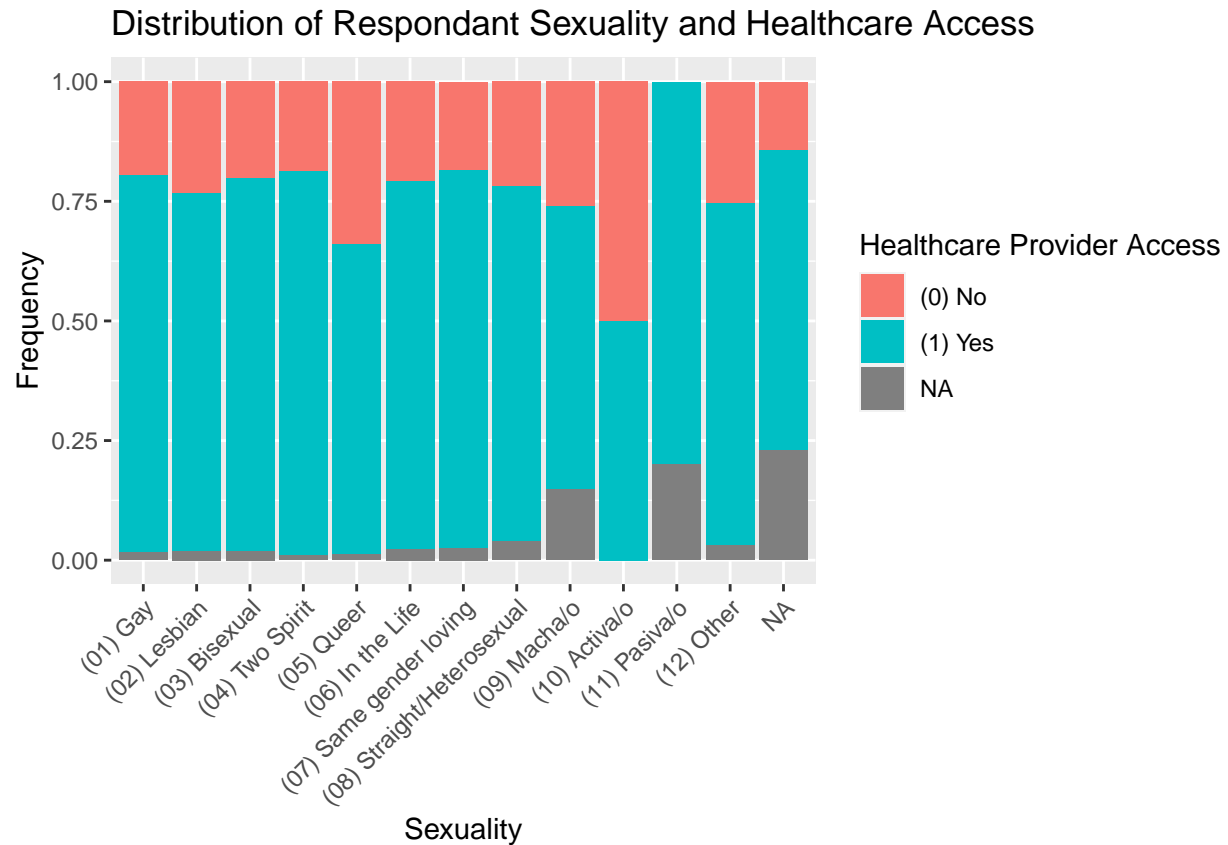
```
#geom_label_repel(nudge_y = 1, direction = "y", hjust = 1) #for if label needs adjusting?
```

This pie chart depicts respondents' gender identities. The biggest group of respondents are male, followed by female, followed by "other." There are more M2F transgender individuals than F2M transgender individuals.

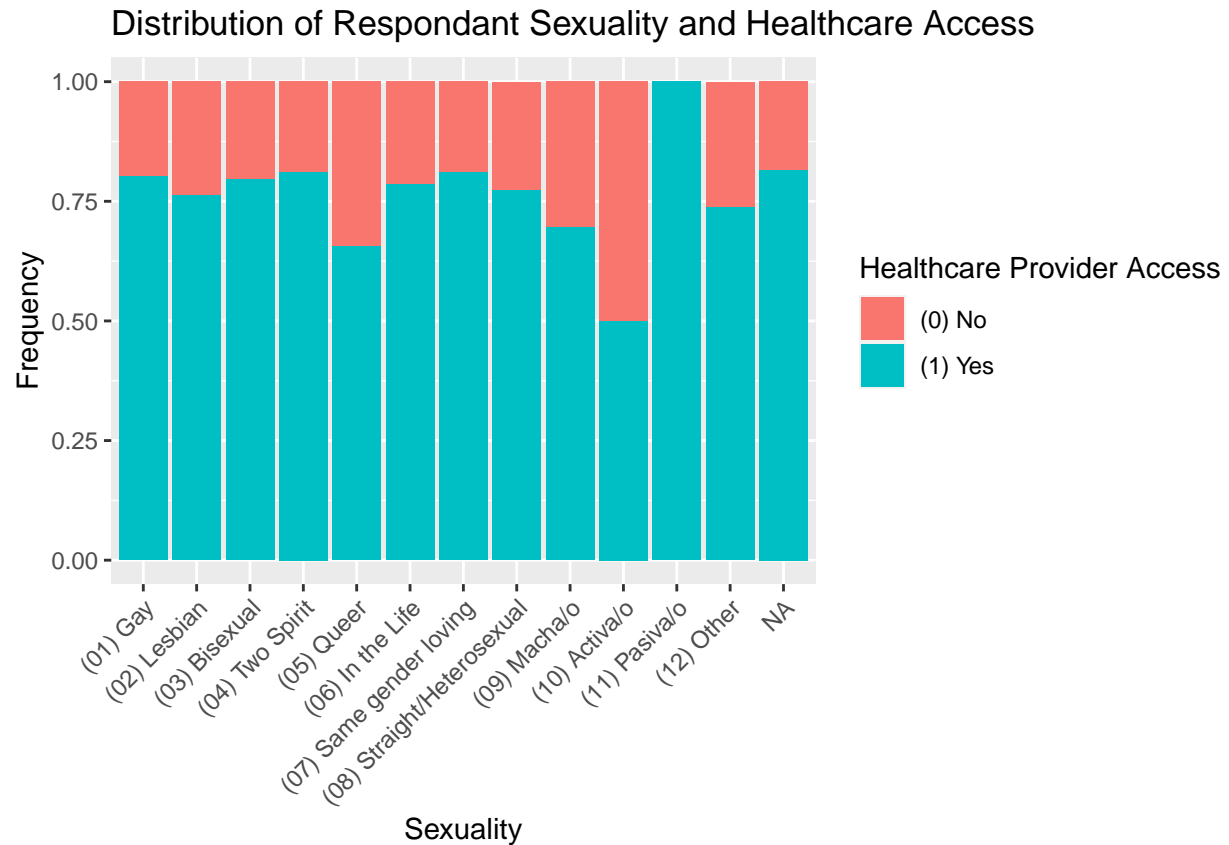
```
ggplot(data, aes(x = sexuality,
                  fill = healthInsureAcc)) +
  geom_bar(position = "fill") +
  labs (x = "Sexuality",
        y = "Frequency",
        fill = "Healthcare Insurance Access",
        title = "Distribution of Respondant Sexuality and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```



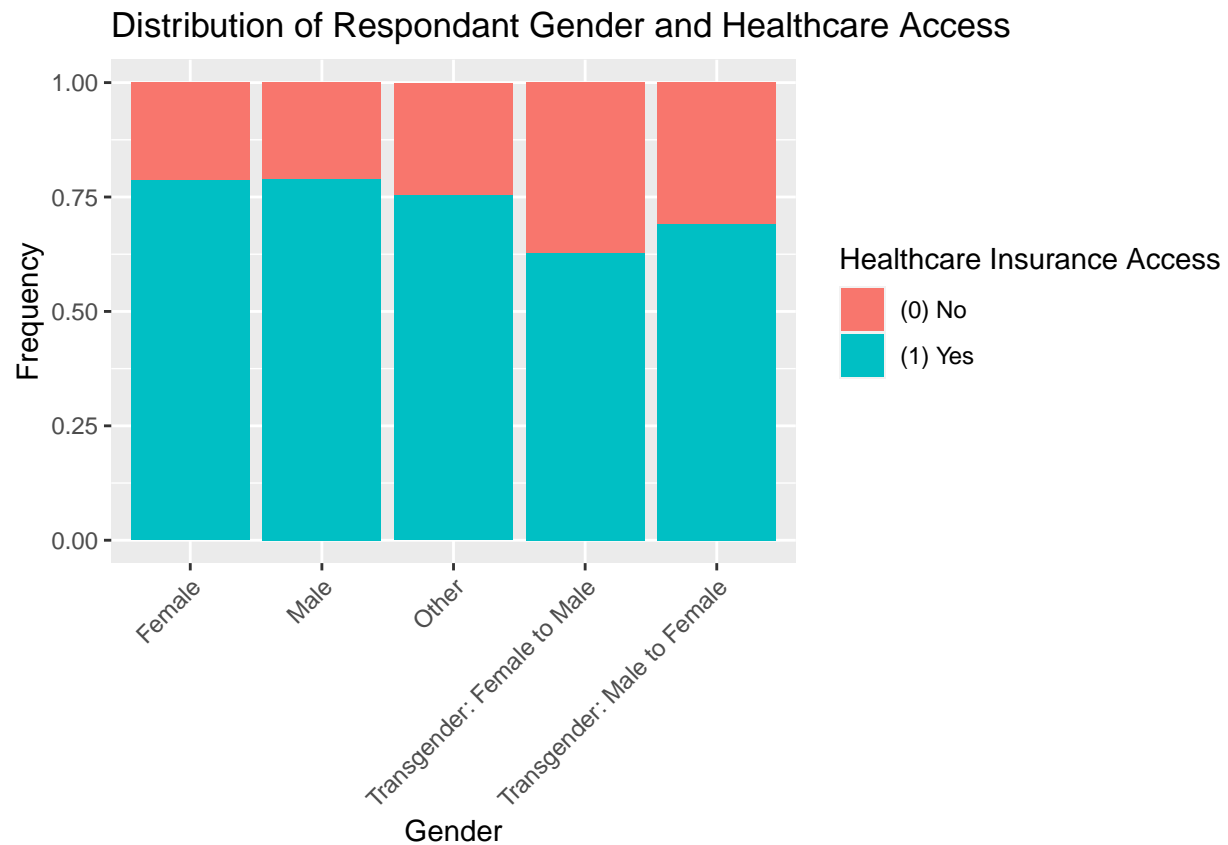
```
data %>%
  ggplot(aes(x = sexuality,
              fill = healthProvideAcc)) +
  geom_bar(position = "fill") +
  labs (x = "Sexuality",
        y = "Frequency",
        fill = "Healthcare Provider Access",
        title = "Distribution of Respondant Sexuality and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```

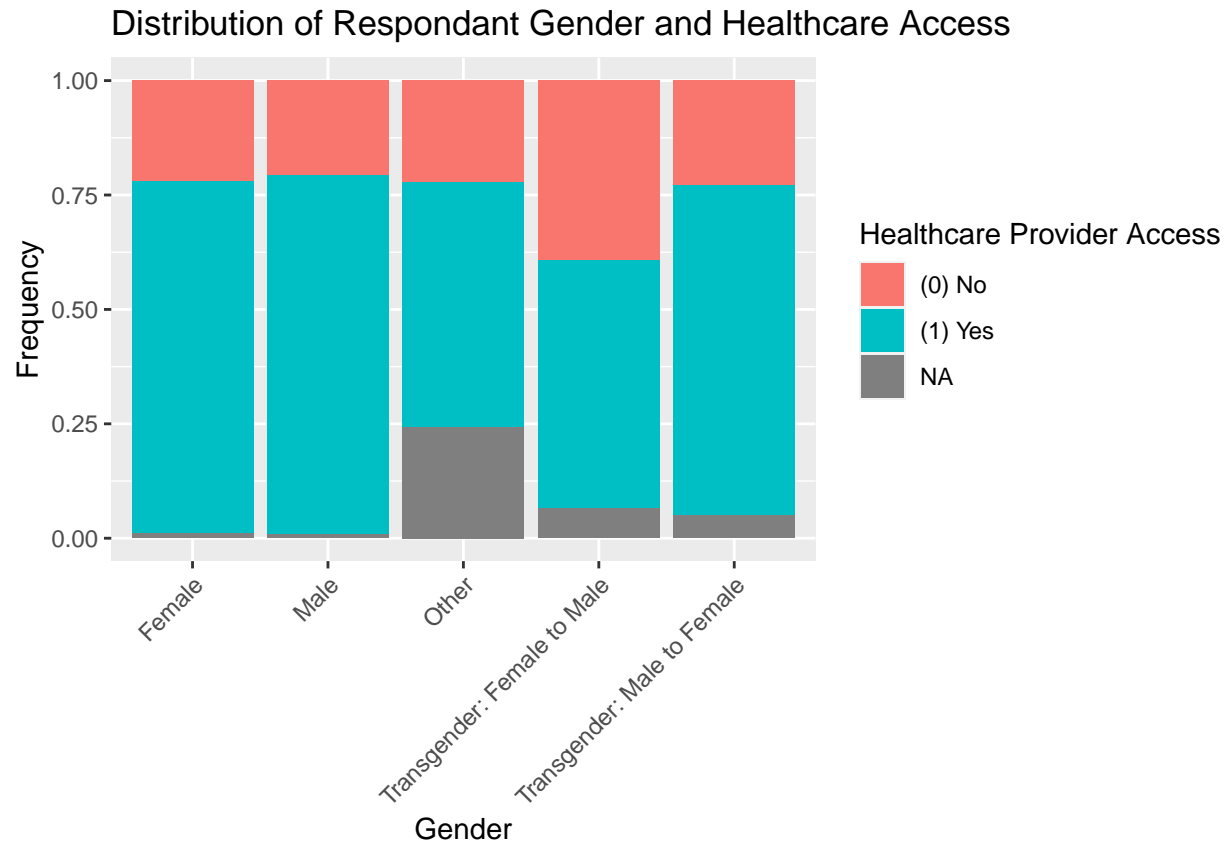
```
data %>%
  filter(is.na(healthProvideAcc) == FALSE) %>%
  ggplot(aes(x = sexuality,
             fill = healthProvideAcc)) +
  geom_bar(position = "fill") +
  labs (x = "Sexuality",
        y = "Frequency",
        fill = "Healthcare Provider Access",
        title = "Distribution of Respondant Sexuality and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```



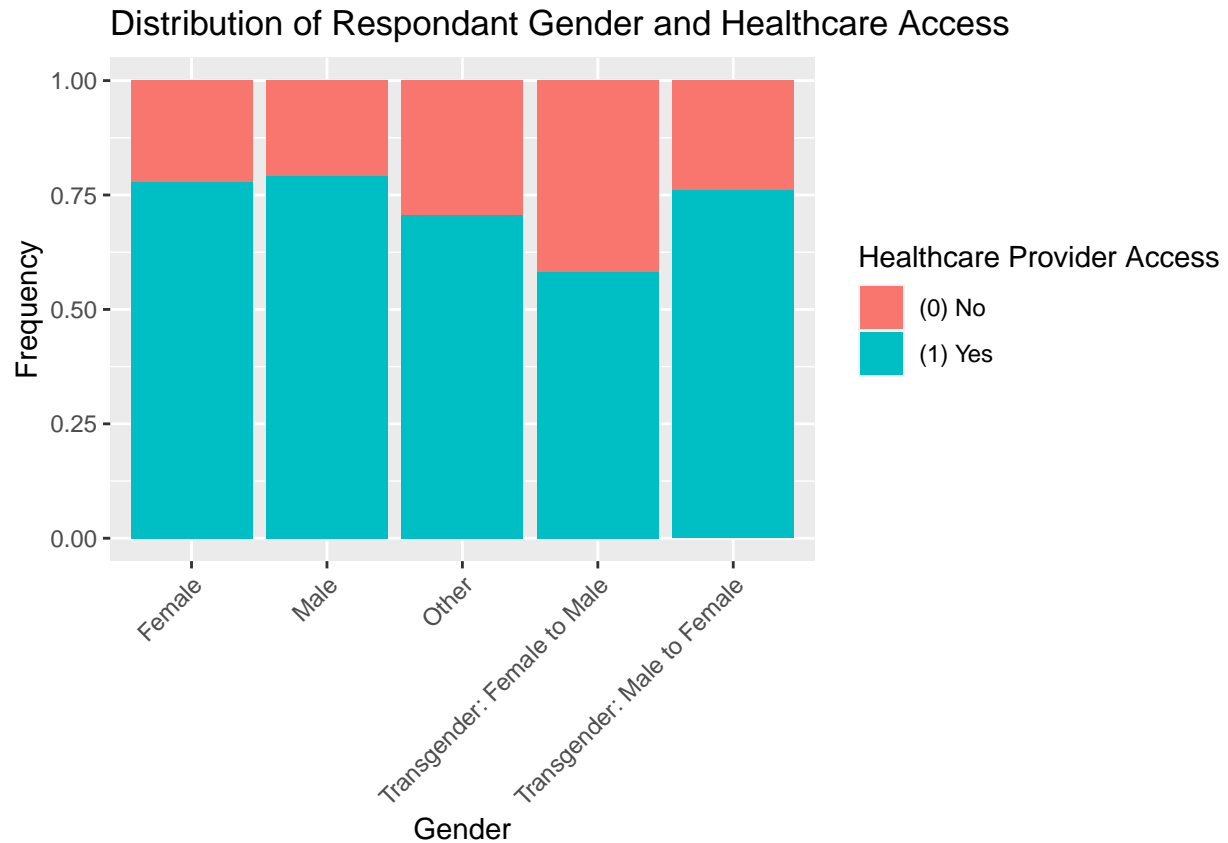
```
data %>%
  filter(is.na(healthInsureAcc) == FALSE) %>%
  ggplot(aes(x = gender,
              fill = healthInsureAcc)) +
  geom_bar(position = "fill") +
  labs (x = "Gender",
        y = "Frequency",
        fill = "Healthcare Insurance Access",
        title = "Distribution of Respondant Gender and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```



```
ggplot(data, aes(x = gender,
                  fill = healthProvideAcc)) +
  geom_bar(position = "fill") +
  labs (x = "Gender",
        y = "Frequency",
        fill = "Healthcare Provider Access",
        title = "Distribution of Respondant Gender and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```



```
#filtered
data %>%
  filter(is.na(healthProvideAcc) == FALSE) %>%
  ggplot(aes(x = gender,
              fill = healthProvideAcc)) +
  geom_bar(position = "fill") +
  labs(x = "Gender",
       y = "Frequency",
       fill = "Healthcare Provider Access",
       title = "Distribution of Respondant Gender and Healthcare Access") +
  theme(axis.text.x = element_text(angle = 45,
                                    hjust = 1))
```



These are responses according to gender when asked, “Do you have a regular doctor or health care provider?” The visualization shows that a greater percentage of transgender individuals answered “no” than cisgender individuals. Also, note that there are more missing values for transgendered individuals than cisgendered respondents, which may provide more insight to the disparity.