U.S. Transgender Survey

Aishwarya

February 28, 2019

Transform the data to include only 3 gender categories for trans men, trans women, and non-binary participants. Use the following definitions when transforming the dataset: (1) trans women are women who were assignedmale-at-birth; (2) trans men are men who were assigned-female-at-birth; (3) combine the "Genderqueer" and "Androgynous" categories to create a single "Non-binary" category. Filter the dataset to include only participants in these categorieS

```
load("C:/Users/Aishwarya/Desktop/NEU/Introduction to Data management/hw4/ICPSR_31721/DS0001/31721-0001-

library(dplyr)
```

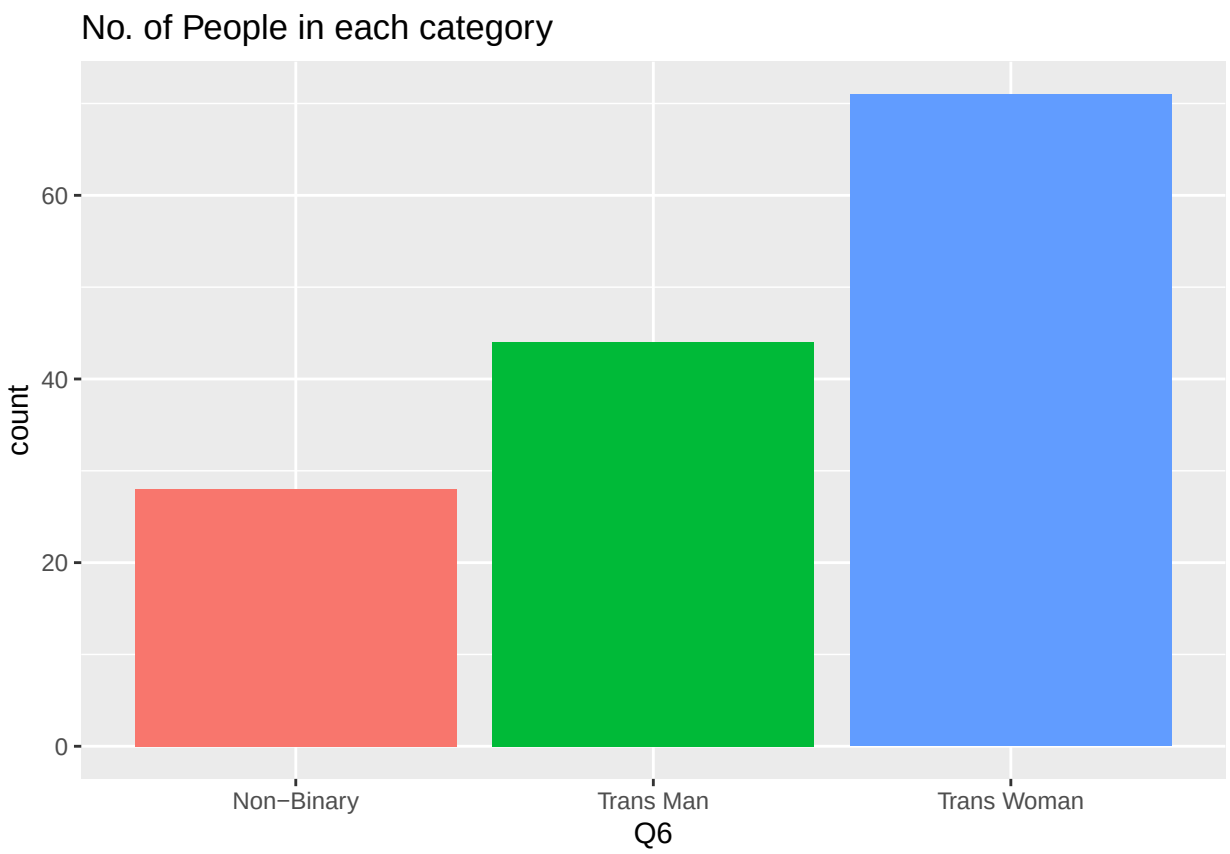
```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)

transwoman <-
  da31721.0001 %>%
  filter(Q5 == "(1) Male" & Q6 == "(2) Woman") %>%
  mutate(Q6 = "Trans Woman")

transman <-
  da31721.0001 %>%
  filter(Q5 == "(2) Female" & Q6 == "(1) Man") %>%
  mutate(Q6 = "Trans Man")

nonbinary <-
  da31721.0001 %>%
  filter(Q6 =="(4) Androgynous" | Q6 == "(6) Gender Queer") %>%
  mutate(Q6 = "Non-Binary")


final <-
  transwoman %>%
  full_join(transman) %>%
  full_join(nonbinary)
```

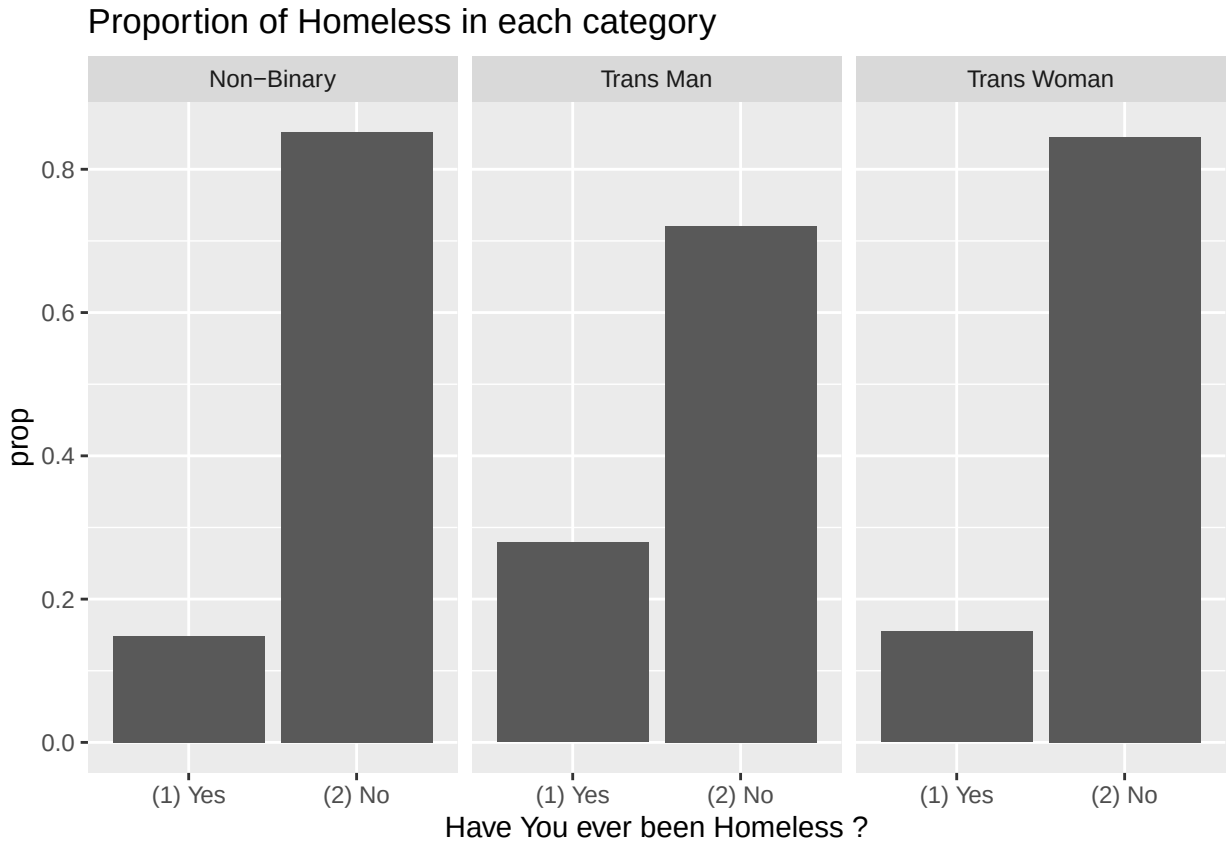
```
## Joining, by = c("RESPKEY", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", "QO6", "Q7", "Q8", "QO8", "D9_1", "D9
```

```
## Joining, by = c("RESPKEY", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", "QO6", "Q7", "Q8", "QO8", "D9_1", "D9_
```

```
ggplot(final) +
  geom_bar(aes(x = Q6, fill = Q6), show.legend = FALSE) +
  ggtitle("No. of People in each category")
```




```
final %>%
  filter(is.na(Q88) == FALSE) %>%
  ggplot(aes(x = Q88)) +
  geom_bar(aes(y= ..prop.., group = 1)) +
  facet_wrap(~Q6) +
  ggtitle("Proportion of Homeless in each category") +
  xlab("Have You ever been Homeless ?")
```




```
#Finding homeless proportion in this survey (not asked in homework)
sum(da31721.0001$Q88 == '(1) Yes' , na.rm = TRUE)/ sum(da31721.0001$Q88 %in% c('(1) Yes', '(2) No'))
```

```
## [1] 0.247093
```

From plot, around 18% of Non- Binary,around 28 % of Trans Man and around 18% of Trans Woman are homeless according to Virginia transgender survey.From calculated result, 24% of respondents in this survey are homeless. But only 6.2% of general population is homeless. Assuming sample space is equal for both surveys, we can possibly say that transgenders are being homeless compared to general population in which transman shares high proportion

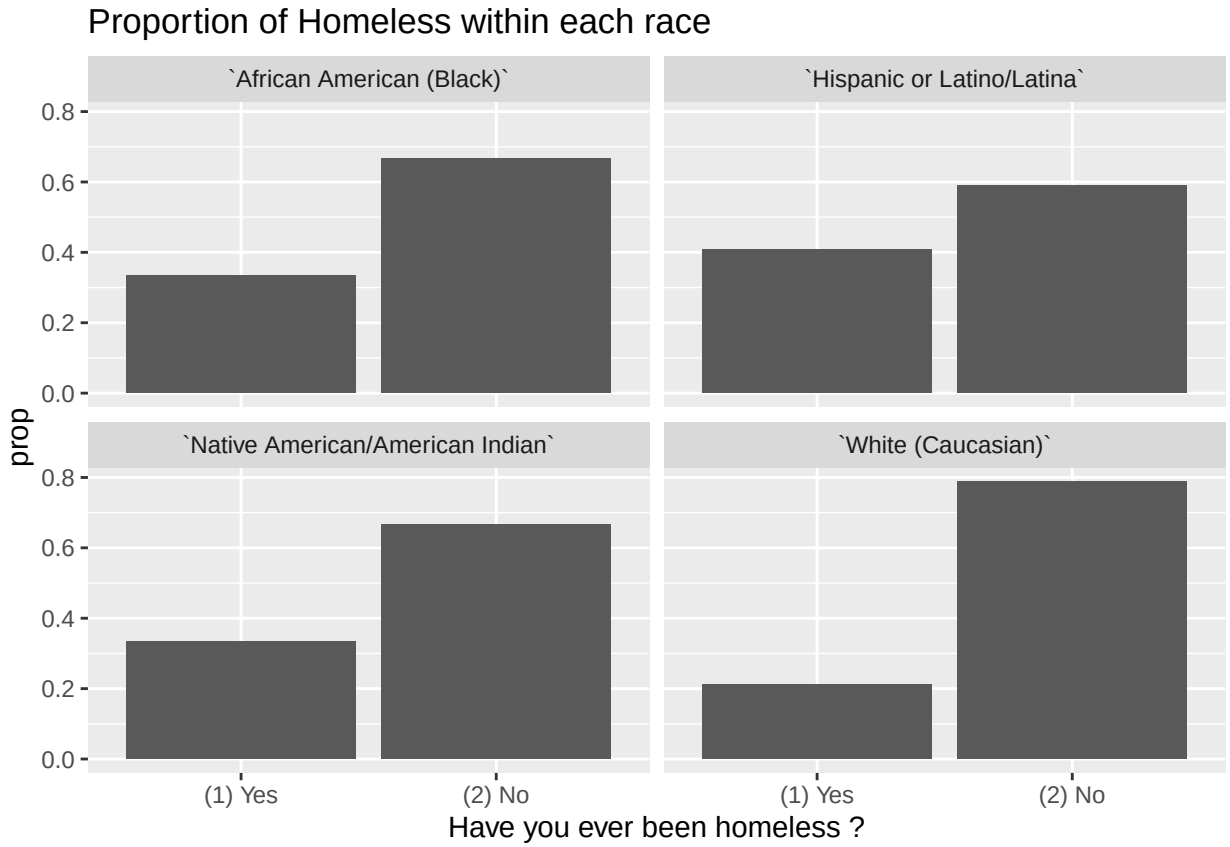
Transform the dataset to have a column for race indicating the race of the participant. Include only the racial demographics with publicly available data (i.e., African American, Caucasian, Hispanic/Latinx, and Native American).

```
`African American (Black)` <-
  da31721.0001 %>%
  filter(D9_1 == "(1) Selected") %>%
  mutate(race = '`African American (Black)`')

`White (Caucasian)` <-
  da31721.0001 %>%
  filter( D9_2 =="(1) Selected" ) %>%
  mutate(race = '`White (Caucasian)`')
```

```r
`Hispanic or Latino/Latina` <-
  da31721.0001 %>%
  filter( D9_3 =="(1) Selected" ) %>%
  mutate(race = '`Hispanic or Latino/Latina`')

`Native American/American Indian` <-
  da31721.0001 %>%
  filter( D9_4 =="(1) Selected" ) %>%
  mutate(race = '`Native American/American Indian`')

Race <- `African American (Black)` %>%
  full_join(`White (Caucasian)`) %>%
  full_join(`Hispanic or Latino/Latina`) %>%
  full_join(`Native American/American Indian`)
```

```
## Joining, by = c("RESPKEY", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", "QO6", "Q7", "Q8", "QO8", "D9_1", "D9
## Joining, by = c("RESPKEY", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", "QO6", "Q7", "Q8", "QO8", "D9_1", "D9
## Joining, by = c("RESPKEY", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", "QO6", "Q7", "Q8", "QO8", "D9_1", "D9
```

```r
Race %>%
  filter(is.na(Q88) == FALSE) %>%
  ggplot(aes(x = Q88)) +
  geom_bar(aes(y = ..prop.., group = 1)) +
  facet_wrap(~race) +
  ggtitle("Proportion of Homeless within each race")+
  xlab("Have you ever been homeless ?")
```




```
#Finding homeless proportion in this survey (not asked in homework)
sum(da31721.0001$Q88 == '(1) Yes' , na.rm = TRUE)/ sum(da31721.0001$Q88 %in% c('(1) Yes', '(2) No'))
```

```
## [1] 0.247093
```

From plot, around 30% of Blacks, around 40% of Hispanic, around 30% of Native Americans and around 20% of whites are homeless according to Virginia transgender survey. From calculated result, 24% of respondents are homeless. But only 6.2% of general population is homeless. Assuming sample space is equal for both surveys, we can possibly say that transgenders are being homeless compared to general population in which Hispanic shares highest proportion.

One of the findings reported in the 2015 U.S. Transgender Survey (http://www.ustranssurvey.org) was that a staggering 40% of the respondents reported attempting suicide in their lifetime, nearly nine times the attempted suicide rate of the general U.S. population (4.6%). calculate the total proportion of participants who have attempted suicide in the Virginia THIS survey.

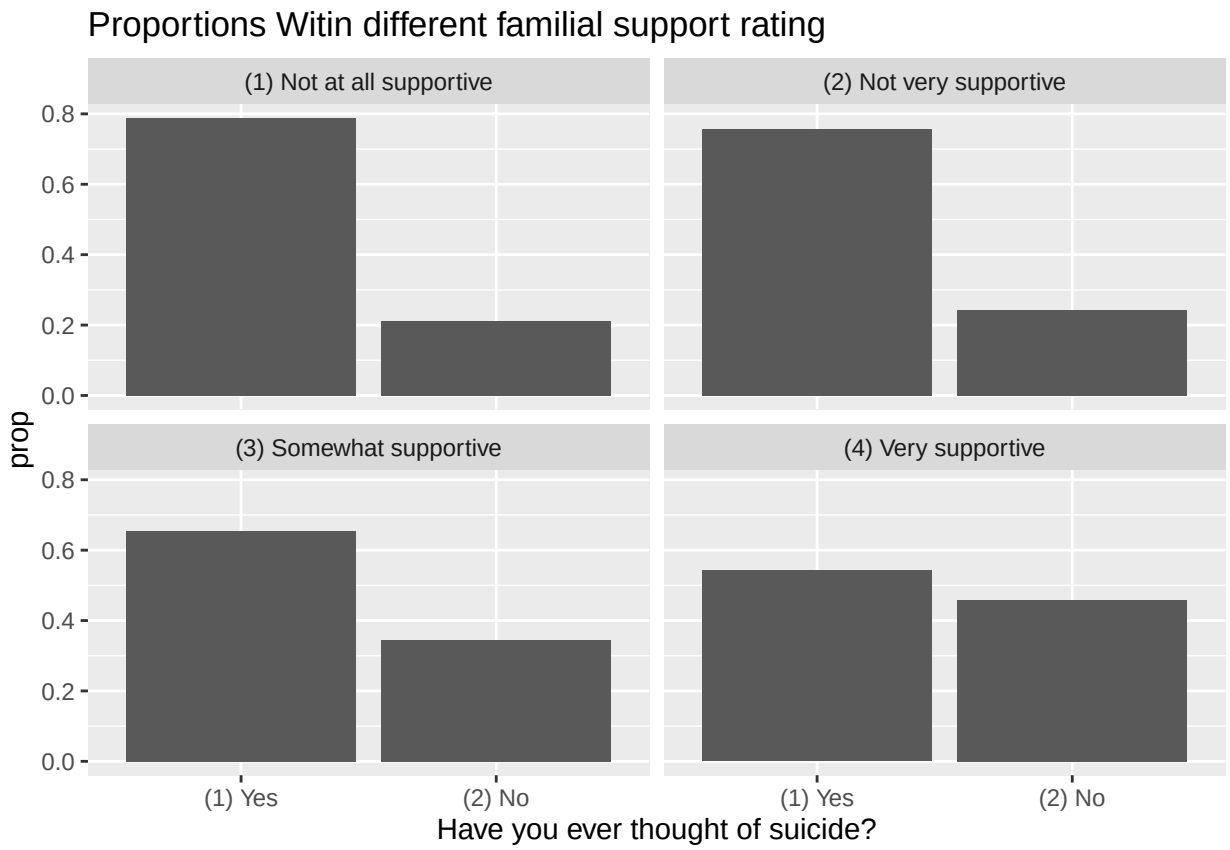
```
#Proportion of participants who have attempted suicide
sum(da31721.0001$Q133 == '(1) Yes',na.rm = TRUE)/
  sum(da31721.0001$Q133 %in% c('(1) Yes', '(2) No', NA))
```

```
## [1] 0.2542857
```

25% of respondents attempted suicide once in their life time which is LOWER than national average of U.S transgender and HIGHER than U.S general population who attempted suicides.

We would like to know if having a birth family who is supportive of one's gender identity and expression reduces the risk of suicide. Using the full dataset, filter the dataset to remove participants who answered "Not applicable to me" to the question about familial support, and then create bar plots showing the proportions of participants who have thought about killing themselves for each level of familial support

```
da31721.0001 %>%
  filter(!(Q119 == '(5) Not applicable to me'))%>%
  #filter(Q131 == '(1) Yes') %>%
  filter(is.na(Q131) == FALSE)%>%
  ggplot(aes(x = Q131)) +
  geom_bar(aes(y = ..prop.., group = 1)) +
  facet_wrap(~Q119) +
  xlab("Have you ever thought of suicide?") +
  ggtitle("Proportions Witin different familial support rating")
```



There is a decrease in suicide rate when the family supportivity increases.