Analyzing Sociodemographic Inequities in Smartphone Usage and Privacy Behaviors

Chriss Jordan Oboa

2024-12-02

#Libraries

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- 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
```

```
library(dplyr)
library(ggplot2)
library(broom)
library(tinytex)
```

#Step 0: Data Inspection

```
# Load the dataset (adjust the path to where your dataset is stored)
Data_Original <- read.csv("~/Downloads/DataAndSociety_DataAccessGrant_SurveyDataFiles-2/



#duplicate the data
Data <- Data_Original

# Inspecting the structure of the dataset
str(Data)
```

```
## 'data.frame':    3000 obs. of  168 variables:
##  $ psraid    : int  100010 100017 100021 100023 100031 100037 100038 100041 100051 10
##  $ sample    : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ int_date  : int  120415 111815 111915 111815 111815 111815 111915 120415 111815 11
##  $ lang      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ comp      : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ version   : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ area      : int  216 716 843 518 718 252 212 903 713 781 ...
##  $ scregion  : int  2 1 3 1 1 3 1 3 3 1 ...
##  $ sfips     : int  39035 36013 45041 36019 36047 37147 36061 48499 48201 25023 ...
##  $ sstate    : int  39 36 45 36 36 37 36 48 48 25 ...
##  $ smsa      : int  17460 0 22500 0 35620 24780 35620 0 26420 14460 ...
##  $ susr      : chr  "S" "R" "S" "R" ...
##  $ sdma      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ sstrata   : int  6 7 8 7 10 9 10 7 8 2 ...
##  $ fips      : int  39035 36013 45041 36019 36047 37195 36061 48499 48201 25025 ...
##  $ usr       : chr  "S" "R" "S" "R" ...
##  $ cregion   : int  2 1 3 1 1 3 1 3 3 1 ...
##  $ state     : int  39 36 45 36 36 37 36 48 48 25 ...
##  $ density   : int  5 2 2 1 5 2 5 1 5 5 ...
##  $ llitext1  : int  1 2 2 2 2 1 1 1 2 2 ...
##  $ qs1       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ sex       : int  2 1 1 2 2 1 2 2 2 1 ...
##  $ q1        : int  4 3 4 2 4 2 2 2 3 1 ...
##  $ q2        : int  3 1 4 1 3 2 1 3 4 4 ...
##  $ q3a       : int  5 3 5 4 2 4 2 3 1 2 ...
##  $ q3b       : int  1 2 8 2 1 1 3 1 1 4 ...
##  $ q3c       : int  3 1 8 2 2 2 2 1 1 4 ...
##  $ q3d       : int  1 3 8 4 1 3 3 4 3 4 ...
##  $ q3e       : int  1 2 5 2 2 4 6 1 1 4 ...
##  $ q3f       : int  1 3 4 3 1 3 2 3 1 4 ...
##  $ q3g       : int  1 3 8 2 4 3 1 1 1 4 ...
##  $ q3h       : int  3 1 2 2 2 2 1 3 1 4 ...
##  $ eminuse   : int  2 1 2 1 1 1 1 1 1 1 ...
##  $ intmob    : int  2 1 2 1 2 1 1 2 2 1 ...
##  $ iuser     : int  0 1 0 1 1 1 1 1 1 1 ...
##  $ home3nw   : int  NA 1 NA 1 1 1 1 1 1 1 ...
##  $ bbhome1   : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ freeint   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ device1a  : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ smart1    : int  2 1 2 1 1 1 1 2 2 1 ...
##  $ q4        : int  NA 1 NA 2 2 2 3 NA NA 2 ...
##  $ q5a       : int  NA 1 NA 1 1 1 1 1 2 1 ...
##  $ q5b       : int  NA 2 NA 2 2 2 1 2 1 2 ...
##  $ q5c       : int  NA 2 NA 2 2 2 2 1 2 2 ...
```

```
##  $ q5d       : int  NA 2 NA 2 2 2 2 2 2 1 ...
##  $ q5e       : int  NA 2 NA 1 2 2 1 1 2 1 ...
##  $ q5f       : int  NA 1 NA 1 1 1 1 2 2 1 ...
##  $ q6        : int  NA 2 NA 8 1 2 1 2 NA 2 ...
##  $ q7        : int  NA 1 NA 1 1 2 2 2 NA 2 ...
##  $ q8        : int  NA NA NA NA NA NA 2 NA 2 NA ...
##  $ q9        : int  NA NA NA NA NA NA NA 3 NA NA ...
##  $ q10a      : int  NA 1 NA 1 3 1 1 5 1 2 ...
##  $ q10b      : int  NA 8 NA 1 3 3 2 3 2 1 ...
##  $ q10c      : int  NA 1 NA 3 3 3 1 3 3 2 ...
##  $ q10d      : int  NA 1 NA 1 3 1 1 5 3 2 ...
##  $ q10e      : int  NA 1 NA 1 3 1 1 3 3 1 ...
##  $ q10f      : int  NA 1 NA 1 1 1 1 2 2 1 ...
##  $ q12a      : int  NA 1 NA 1 1 1 1 1 2 1 ...
##  $ q12b      : int  NA 1 NA 1 2 1 1 2 2 1 ...
##  $ q12c      : int  NA 1 NA 1 1 2 2 1 2 1 ...
##  $ q12d      : int  NA 1 NA 1 1 2 2 3 2 1 ...
##  $ q12e      : int  NA 1 NA 1 2 1 3 1 2 2 ...
##  $ q12f      : int  NA 1 NA 1 1 1 2 2 2 3 ...
##  $ q12g      : int  NA 1 NA 1 1 1 1 1 2 1 ...
##  $ employ    : int  3 3 3 2 1 1 2 2 3 1 ...
##  $ notw      : int  1 1 4 NA NA NA NA NA 1 NA ...
##  $ par       : int  2 2 2 1 2 1 1 2 2 2 ...
##  $ q13a      : int  3 3 4 3 3 2 4 1 2 1 ...
##  $ q13b      : int  3 2 4 1 4 1 4 8 2 1 ...
##  $ q13c      : int  3 2 4 2 4 2 4 1 2 1 ...
##  $ q13d      : int  1 2 4 2 4 2 4 1 1 1 ...
##  $ q13e      : int  NA 2 NA 2 3 3 4 5 2 1 ...
##  $ q13f      : int  NA 2 NA 3 2 3 4 1 1 1 ...
##  $ q13g      : int  NA 2 NA 2 3 3 4 5 3 1 ...
##  $ q13h      : int  NA 2 NA 3 3 2 4 8 NA 4 ...
##  $ q13i      : int  NA NA NA 1 3 1 1 5 NA 1 ...
##  $ q13j      : int  NA NA NA 3 NA 2 1 NA NA NA ...
##  $ q14       : int  3 2 3 3 2 2 3 1 2 4 ...
##  $ q15a      : int  NA 2 NA 2 2 2 2 2 2 1 ...
##  $ q15b      : int  NA 2 NA 2 2 1 1 2 1 2 ...
##  $ q15c      : int  NA 2 NA 2 1 2 2 2 2 1 ...
##  $ q15d      : int  NA 2 NA 2 2 2 1 2 2 2 ...
##  $ q15e      : int  NA 2 NA 2 2 2 1 8 2 2 ...
##  $ q15f      : int  NA 2 NA 2 2 2 1 8 2 2 ...
##  $ q15g      : int  NA 1 NA 2 1 2 2 3 2 1 ...
##  $ q15h      : int  NA 2 NA 2 2 1 1 2 2 1 ...
##  $ q15i      : int  NA 2 NA 2 2 1 1 1 1 1 ...
##  $ q15j      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q15x      : int  NA 2 NA 3 3 2 3 2 2 4 ...
```

```
##  $ q16a      : int  2 1 2 1 2 2 2 2 1 2 ...
##  $ q16b      : int  2 1 2 2 2 2 8 2 2 2 ...
##  $ q16c      : int  8 2 2 2 2 1 8 8 1 2 ...
##  $ q16d      : int  NA 2 NA 2 2 2 1 2 1 2 ...
##  $ q16e      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q16f      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q16g      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q16h      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q16i      : int  NA 2 NA 2 2 2 2 2 2 2 ...
##  $ q16j      : int  NA 2 NA 1 2 8 2 8 2 1 ...
##   [list output truncated]
```

```
# Viewing the first few rows
head(Data)
```

```
##   psraid sample int_date lang comp version area scregion sfips sstate  smsa
## 1 100010      1   120415    1    1       1  216        2 39035     39 17460
## 2 100017      1   111815    1    1       1  716        1 36013     36     0
## 3 100021      1   111915    1    1       1  843        3 45041     45 22500
## 4 100023      1   111815    1    1       1  518        1 36019     36     0
## 5 100031      1   111815    1    1       1  718        1 36047     36 35620
## 6 100037      1   111815    1    1       1  252        3 37147     37 24780
##   susr sdma sstrata  fips usr cregion state density llitext1 qs1 sex q1 q2 q3a
## 1    S   NA       6 39035   S       2    39       5        1  NA   2  4  3   5
## 2    R   NA       7 36013   R       1    36       2        2  NA   1  3  1   3
## 3    S   NA       8 45041   S       3    45       2        2  NA   1  4  4   5
## 4    R   NA       7 36019   R       1    36       1        2  NA   2  2  1   4
## 5    U   NA      10 36047   U       1    36       5        2  NA   2  4  3   2
## 6    S   NA       9 37195   S       3    37       2        1  NA   1  2  2   4
##   q3b q3c q3d q3e q3f q3g q3h eminuse intmob iuser home3nw bbhome1 freeint
## 1   1   3   1   1   1   1   3       2      2     0      NA      NA       NA
## 2   2   1   3   2   3   3   1       1      1     1       1       2       NA
## 3   8   8   8   5   4   8   2       2      2     0      NA      NA       NA
## 4   2   2   4   2   3   2   2       1      1     1       1       2       NA
## 5   1   2   1   2   1   4   2       1      2     1       1       2       NA
## 6   1   2   3   4   3   3   2       1      1     1       1       2       NA
##   device1a smart1 q4 q5a q5b q5c q5d q5e q5f q6 q7 q8 q9 q10a q10b q10c q10d
## 1        1      2 NA  NA  NA  NA  NA  NA  NA NA NA NA NA   NA   NA   NA   NA
## 2        1      1  1   1   2   2   2   2   1  2  1 NA NA    1    8    1    1
## 3        1      2 NA  NA  NA  NA  NA  NA  NA NA NA NA NA   NA   NA   NA   NA
## 4        1      1  2   1   2   2   2   1   1  8  1 NA NA    1    1    3    1
## 5        1      1  2   1   2   2   2   2   1  1  1 NA NA    3    3    3    3
## 6        1      1  2   1   2   2   2   2   1  2  2 NA NA    1    3    3    1
##   q10e q10f q12a q12b q12c q12d q12e q12f q12g employ notw par q13a q13b q13c
```

```
## 1   NA   NA   NA   NA   NA   NA   NA   NA   NA      3    1   2    3    3    3
## 2    1    1    1    1    1    1    1    1    1      3    1   2    3    2    2
## 3   NA   NA   NA   NA   NA   NA   NA   NA   NA      3    4   2    4    4    4
## 4    1    1    1    1    1    1    1    1    1      2   NA   1    3    1    2
## 5    3    1    1    2    1    1    2    1    1      1   NA   2    3    4    4
## 6    1    1    1    1    2    2    1    1    1      1   NA   1    2    1    2
##   q13d q13e q13f q13g q13h q13i q13j q14 q15a q15b q15c q15d q15e q15f q15g
## 1    1   NA   NA   NA   NA   NA   NA   3   NA   NA   NA   NA   NA   NA   NA
## 2    2    2    2    2    2   NA   NA   2    2    2    2    2    2    2    1
## 3    4   NA   NA   NA   NA   NA   NA   3   NA   NA   NA   NA   NA   NA   NA
## 4    2    2    3    2    3    1    3   3    2    2    2    2    2    2    2
## 5    4    3    2    3    3    3   NA   2    2    2    1    2    2    2    1
## 6    2    3    3    3    2    1    2   2    2    1    2    2    2    2    2
##   q15h q15i q15j q15x q16a q16b q16c q16d q16e q16f q16g q16h q16i q16j q17a
## 1   NA   NA   NA   NA    2    2    8   NA   NA   NA   NA   NA   NA   NA   NA
## 2    2    2    2    2    1    1    2    2    2    2    2    2    2    2    2
## 3   NA   NA   NA   NA    2    2    2   NA   NA   NA   NA   NA   NA   NA   NA
## 4    2    2    2    3    1    2    2    2    2    2    2    2    2    1    2
## 5    2    2    2    3    2    2    2    2    2    2    2    2    2    2    2
## 6    1    1    2    2    2    2    1    2    2    2    2    2    2    8    1
##   q17b q17c q17d q17e q17f q17g q17h sm1a sm1b sm1c sm1d module occup1 q18a
## 1   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA      9     NA   NA
## 2    2    2    2    2    2    2    2    1    2    2    1      9     NA   NA
## 3   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA      9     NA   NA
## 4    2    2    2    2    2    2    2    1    2    2    2      2      7    2
## 5    2    2    2    2    2    2    2    2    2    1    2      2      1    1
## 6    2    2    2    2    2    2    2    1    1    1    2      2      1    1
##   q18b q18c q19 q20a q20b q20c q20d q20e q20f q20g q21a q21b q21c q22a q22b
## 1   NA   NA  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 2   NA   NA  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 3   NA   NA  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 4    2    2  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 5    2    8   1   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
## 6    1    1   2   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
##   q22c q22d q22e ql1a qc1 age marital hh1 educ2 vet1 disa public party partyln
## 1   NA   NA   NA   NA  NA  84       5   2     3    2    2      1     2      NA
## 2   NA   NA   NA   NA  NA  99       1   2     6    1    1      1     9       9
## 3   NA   NA   NA   NA  NA  47       1   2     1    2    1      1     4       2
## 4   NA   NA   NA   NA  NA  45       1   2     6    2    2      1     3       1
## 5   NA   NA   NA   NA  NA  59       5   1     6    2    2      2     2      NA
## 6   NA   NA   NA   NA  NA  36       1   2     6    2    2      2     1      NA
##   hisp racem1 racem2 racem3 racem4 birth_hisp inc inc1 verify q17hos occup1os
## 1    2      2     NA     NA     NA         NA   2   NA      1
## 2    9      9     NA     NA     NA         NA   9   NA      1
## 3    2      1     NA     NA     NA         NA   1   NA      1
```

```
## 4    2      1     NA     NA     NA         NA   7   NA      1
## 5    2      2     NA     NA     NA         NA   6   NA      1
## 6    2      1     NA     NA     NA         NA   6   NA      1
##   raceos recage2 recincome racecmb racethn racethn2 receduc i_user SmartPhone
## 1              4         1       2       2        2       2      0          2
## 2              9         2       9       9        9       4      1          1
## 3              2         1       1       1        1       1      0          2
## 4              2         2       1       1        1       4      1          1
## 5              3         2       2       2        2       4      1          1
## 6              2         2       1       1        1       4      1          1
##   partysum iphoneuse hphoneuse    weight   standwt
## 1        2         2         2 10.161245 1.1823650
## 2        9         2         2  4.381396 0.5098204
## 3        2         2         2 11.436218 1.3307213
## 4        1         2         2  6.993012 0.8137086
## 5        2         2         2  1.000000 0.1163603
## 6        1         2         2  2.837643 0.3301889
```

```
# Checking for summary statistics and missing values
summary(Data)
```

```
##      psraid           sample        int_date          lang            comp
##  Min.   :100010   Min.   :1.00   Min.   :111815   Min.   :1.000   Min.   :1
##  1st Qu.:104108   1st Qu.:1.00   1st Qu.:120315   1st Qu.:1.000   1st Qu.:1
##  Median :202590   Median :2.00   Median :121415   Median :1.000   Median :1
##  Mean   :169998   Mean   :1.65   Mean   :119567   Mean   :1.086   Mean   :1
##  3rd Qu.:207402   3rd Qu.:2.00   3rd Qu.:121915   3rd Qu.:1.000   3rd Qu.:1
##  Max.   :213234   Max.   :2.00   Max.   :122315   Max.   :2.000   Max.   :1
##
##     version       area          scregion         sfips           sstate
##  Min.   :1   Min.   :201.0   Min.   :1.000   Min.   : 1005   Min.   : 1.00
##  1st Qu.:1   1st Qu.:337.0   1st Qu.:2.000   1st Qu.:11001   1st Qu.:11.00
##  Median :1   Median :575.0   Median :3.000   Median :25025   Median :25.00
##  Mean   :1   Mean   :578.2   Mean   :2.855   Mean   :26434   Mean   :26.34
##  3rd Qu.:1   3rd Qu.:806.0   3rd Qu.:4.000   3rd Qu.:42043   3rd Qu.:42.00
##  Max.   :1   Max.   :989.0   Max.   :4.000   Max.   :56043   Max.   :56.00
##
##       smsa           susr                sdma          sstrata
##  Min.   :    0   Length:3000        Min.   :500.0   Min.   : 1.000
##  1st Qu.:12260   Class :character   1st Qu.:530.0   1st Qu.: 7.000
##  Median :26420   Mode  :character   Median :623.0   Median : 8.000
##  Mean   :23417                      Mean   :646.2   Mean   : 7.891
##  3rd Qu.:35620                      3rd Qu.:757.0   3rd Qu.: 9.000
##  Max.   :49740                      Max.   :881.0   Max.   :10.000
```

```
##                                    NA's   :1050
##       fips           usr               cregion          state
##  Min.   : 1003   Length:3000        Min.   :1.000   Min.   : 1.00
##  1st Qu.:12001   Class :character   1st Qu.:2.000   1st Qu.:12.00
##  Median :26052   Mode  :character   Median :3.000   Median :26.00
##  Mean   :26724                      Mean   :2.857   Mean   :26.63
##  3rd Qu.:42084                      3rd Qu.:4.000   3rd Qu.:42.00
##  Max.   :56013                      Max.   :4.000   Max.   :56.00
##
##     density         llitext1         qs1            sex             q1
##  Min.   :1.000   Min.   :1.00   Min.   :2      Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.00   1st Qu.:2      1st Qu.:1.000   1st Qu.:2.000
##  Median :3.000   Median :2.00   Median :2      Median :1.000   Median :3.000
##  Mean   :2.808   Mean   :1.51   Mean   :2      Mean   :1.476   Mean   :2.889
##  3rd Qu.:4.000   3rd Qu.:2.00   3rd Qu.:2      3rd Qu.:2.000   3rd Qu.:4.000
##  Max.   :5.000   Max.   :2.00   Max.   :2      Max.   :2.000   Max.   :9.000
##                  NA's   :1950   NA's   :1050
##        q2             q3a             q3b             q3c             q3d
##  Min.   :1.00   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.00   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :2.00   Median :3.000   Median :2.000   Median :2.000   Median :3.000
##  Mean   :2.36   Mean   :2.788   Mean   :2.313   Mean   :1.914   Mean   :2.847
##  3rd Qu.:3.00   3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:4.000
##  Max.   :9.00   Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##
##       q3e             q3f             q3g             q3h
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :2.000   Median :3.000   Median :3.000   Median :2.000
##  Mean   :2.366   Mean   :2.585   Mean   :2.801   Mean   :2.062
##  3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:3.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##
##     eminuse          intmob          iuser           home3nw
##  Min.   :1.000   Min.   :1.000   Min.   :0.0000   Min.   :1.00
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.0000   1st Qu.:1.00
##  Median :1.000   Median :1.000   Median :1.0000   Median :1.00
##  Mean   :1.253   Mean   :1.352   Mean   :0.7833   Mean   :1.11
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:1.0000   3rd Qu.:1.00
##  Max.   :8.000   Max.   :9.000   Max.   :1.0000   Max.   :8.00
##                                                   NA's   :650
##     bbhome1         freeint         device1a         smart1          q4
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.00   1st Qu.:1.000
##  Median :2.000   Median :2.000   Median :1.000   Median :1.00   Median :2.000
```

```
##  Mean   :2.291   Mean   :1.721   Mean   :1.212   Mean   :1.98   Mean   :1.828
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:2.00   3rd Qu.:2.000
##  Max.   :9.000   Max.   :8.000   Max.   :2.000   Max.   :9.00   Max.   :8.000
##  NA's   :896     NA's   :2756    NA's   :1950    NA's   :223    NA's   :1276
##       q5a             q5b             q5c             q5d            q5e
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.00   1st Qu.:1.000
##  Median :1.000   Median :2.000   Median :2.000   Median :2.00   Median :2.000
##  Mean   :1.325   Mean   :1.576   Mean   :1.816   Mean   :1.89   Mean   :1.541
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.00   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650    NA's   :650
##       q5f             q6              q7              q8
##  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.:2.000
##  Median :1.000   Median :2.000   Median :2.000   Median :2.000
##  Mean   :1.248   Mean   :2.137   Mean   :1.742   Mean   :2.203
##  3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:3.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :8.000
##  NA's   :650     NA's   :1387    NA's   :1387    NA's   :1984
##        q9             q10a            q10b            q10c
##  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 :2.000   Median :1.000   Median :2.000   Median :3.000
##  Mean   :2.142   Mean   :1.788   Mean   :2.212   Mean   :2.395
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :2513    NA's   :650     NA's   :650     NA's   :650
##       q10d            q10e            q10f            q12a
##  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 :2.000   Median :1.000
##  Mean   :2.031   Mean   :2.152   Mean   :2.224   Mean   :1.275
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:1.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650
##       q12b            q12c            q12d            q12e           q12f
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.00   1st Qu.:1.000
##  Median :1.000   Median :1.000   Median :1.000   Median :1.00   Median :1.000
##  Mean   :1.445   Mean   :1.463   Mean   :1.683   Mean   :1.46   Mean   :1.658
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.00   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650    NA's   :650
##       q12g           employ            notw            par
```

```
##  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 :1.000   Median :2.000   Median :1.000   Median :2.000
##  Mean   :1.432   Mean   :2.102   Mean   :2.321   Mean   :1.805
##  3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :650                     NA's   :1598
##       q13a            q13b            q13c            q13d            q13e
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00
##  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.00
##  Median :2.000   Median :2.000   Median :2.000   Median :2.000   Median :3.00
##  Mean   :2.663   Mean   :2.131   Mean   :2.053   Mean   :2.359   Mean   :2.72
##  3rd Qu.:4.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:4.00
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.00
##                                                  NA's   :223     NA's   :650
##       q13f            q13g            q13h            q13i
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :2.000   Median :2.000   Median :3.000   Median :1.000
##  Mean   :2.414   Mean   :2.634   Mean   :2.893   Mean   :1.773
##  3rd Qu.:3.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :1387    NA's   :1418
##       q13j            q14             q15a            q15b
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   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   :2.221   Mean   :2.354   Mean   :1.876   Mean   :2.225
##  3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :2345                    NA's   :650     NA's   :650
##       q15c            q15d            q15e            q15f
##  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.:2.000
##  Median :2.000   Median :2.000   Median :2.000   Median :2.000
##  Mean   :1.893   Mean   :1.853   Mean   :2.021   Mean   :1.909
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :8.000   Max.   :9.000   Max.   :8.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650
##       q15g            q15h            q15i            q15j            q15x
##  Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.00   1st Qu.:2.000   1st Qu.:1.000
##  Median :2.000   Median :1.000   Median :1.00   Median :2.000   Median :2.000
##  Mean   :1.583   Mean   :1.619   Mean   :1.55   Mean   :1.933   Mean   :2.096
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00   3rd Qu.:2.000   3rd Qu.:3.000
```

```
##  Max.   :9.000   Max.   :9.000   Max.   :9.00   Max.   :8.000   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650    NA's   :650     NA's   :650
##       q16a            q16b            q16c           q16d
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   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   :2.085   Mean   :2.487   Mean   :2.437   Mean   :2.099
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##                                                  NA's   :650
##       q16e            q16f            q16g           q16h
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   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   :2.078   Mean   :2.006   Mean   :1.966   Mean   :2.132
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650
##       q16i            q16j            q17a           q17b
##  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.:1.000
##  Median :2.000   Median :2.000   Median :2.000   Median :2.000
##  Mean   :1.951   Mean   :2.147   Mean   :1.631   Mean   :1.632
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :8.000   Max.   :8.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650
##       q17c            q17d            q17e           q17f
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:2.000   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   :1.823   Mean   :1.946   Mean   :1.913   Mean   :1.879
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :8.000   Max.   :8.000   Max.   :8.000   Max.   :9.000
##  NA's   :650     NA's   :650     NA's   :650     NA's   :650
##       q17g            q17h            sm1a           sm1b           sm1c
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.00
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.00
##  Median :2.000   Median :2.000   Median :1.000   Median :2.000   Median :2.00
##  Mean   :1.935   Mean   :2.012   Mean   :1.438   Mean   :1.713   Mean   :1.57
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00
##  Max.   :8.000   Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.00
##  NA's   :650     NA's   :650     NA's   :1217    NA's   :1217    NA's   :1217
##       sm1d           module          occup1          q18a
##  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 :2.000   Median : 4.000   Median :2.000
##  Mean   :1.472   Mean   :4.748   Mean   : 6.103   Mean   :1.903
##  3rd Qu.:2.000   3rd Qu.:9.000   3rd Qu.: 8.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :99.000   Max.   :9.000
##  NA's   :1217                    NA's   :1661     NA's   :1661
##       q18b            q18c            q19             q20a
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000
##  Median :2.000   Median :2.000   Median :2.000   Median :1.000
##  Mean   :1.801   Mean   :1.972   Mean   :2.063   Mean   :1.381
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:1.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##  NA's   :1661    NA's   :1661    NA's   :2286    NA's   :2577
##       q20b            q20c            q20d            q20e
##  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 :1.000   Median :1.000   Median :1.000
##  Mean   :1.395   Mean   :1.473   Mean   :1.452   Mean   :1.362
##  3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1.000
##  Max.   :5.000   Max.   :9.000   Max.   :9.000   Max.   :8.000
##  NA's   :2577    NA's   :2577    NA's   :2577    NA's   :2577
##       q20f            q20g            q21a            q21b           q21c
##  Min.   :1.000   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.:1.000   1st Qu.:1.00
##  Median :1.000   Median :1.000   Median :1.000   Median :1.000   Median :1.00
##  Mean   :1.643   Mean   :1.291   Mean   :1.466   Mean   :1.738   Mean   :1.56
##  3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:1.00
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :8.00
##  NA's   :2577    NA's   :2577    NA's   :2577    NA's   :2577    NA's   :2577
##       q22a            q22b            q22c            q22d
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   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   :1.745   Mean   :1.619   Mean   :1.868   Mean   :1.825
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :3.000   Max.   :9.000   Max.   :8.000   Max.   :8.000
##  NA's   :2577    NA's   :2577    NA's   :2577    NA's   :2577
##       q22e            ql1a            qc1             age            marital
##  Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :18.00   Min.   :1.00
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:36.00   1st Qu.:1.00
##  Median :2.000   Median :2.000   Median :2.000   Median :54.00   Median :3.00
##  Mean   :1.766   Mean   :1.659   Mean   :1.632   Mean   :52.97   Mean   :3.12
##  3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:67.00   3rd Qu.:6.00
##  Max.   :9.000   Max.   :8.000   Max.   :9.000   Max.   :99.00   Max.   :9.00
##  NA's   :2577    NA's   :2777    NA's   :1050
```

```
##       hh1            educ2             vet1            disa
##  Min.   :1.00   Min.   : 1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.00   1st Qu.: 3.000   1st Qu.:2.000   1st Qu.:2.000
##  Median :2.00   Median : 4.000   Median :2.000   Median :2.000
##  Mean   :2.22   Mean   : 5.508   Mean   :1.878   Mean   :1.869
##  3rd Qu.:3.00   3rd Qu.: 6.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :9.00   Max.   :99.000   Max.   :9.000   Max.   :9.000
##
##      public          party          partyln          hisp
##  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   :1.727   Mean   :2.588   Mean   :4.886   Mean   :1.874
##  3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:8.000   3rd Qu.:2.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##                                  NA's   :1611
##      racem1          racem2          racem3      racem4          birth_hisp
##  Min.   :1.000   Min.   :1.000   Min.   :1      Mode:logical   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1      NA's:3000      1st Qu.:1.000
##  Median :1.000   Median :5.000   Median :3                     Median :3.000
##  Mean   :2.137   Mean   :3.639   Mean   :3                     Mean   :2.138
##  3rd Qu.:2.000   3rd Qu.:5.000   3rd Qu.:5                     3rd Qu.:3.000
##  Max.   :9.000   Max.   :6.000   Max.   :5                     Max.   :9.000
##                  NA's   :2928    NA's   :2992                  NA's   :2440
##       inc             inc1           verify     q17hos
##  Min.   : 1.00   Min.   :1.000   Min.   :1   Length:3000
##  1st Qu.: 3.00   1st Qu.:1.000   1st Qu.:1   Class :character
##  Median : 6.00   Median :2.000   Median :1   Mode  :character
##  Mean   :20.51   Mean   :4.925   Mean   :1
##  3rd Qu.: 8.00   3rd Qu.:9.000   3rd Qu.:1
##  Max.   :99.00   Max.   :9.000   Max.   :1
##                  NA's   :2494
##    occup1os            raceos             recage2        recincome
##  Length:3000        Length:3000        Min.   :1.000   Min.   :1.000
##  Class :character   Class :character   1st Qu.:2.000   1st Qu.:1.000
##  Mode  :character   Mode  :character   Median :3.000   Median :2.000
##                                        Mean   :2.893   Mean   :2.098
##                                        3rd Qu.:4.000   3rd Qu.:2.000
##                                        Max.   :9.000   Max.   :9.000
##
##     racecmb          racethn          racethn2        receduc
##  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 :1.000   Median :1.000   Median :1.000   Median :3.000
##  Mean   :2.001   Mean   :1.919   Mean   :2.084   Mean   :2.882
```

```
##  3rd Qu.:2.000   3rd Qu.:3.000   3rd Qu.:3.000   3rd Qu.:4.000
##  Max.   :9.000   Max.   :9.000   Max.   :9.000   Max.   :9.000
##
##      i_user          SmartPhone      partysum       iphoneuse
##  Min.   :0.0000   Min.   :1.00   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.0000   1st Qu.:1.00   1st Qu.:1.000   1st Qu.:2.000
##  Median :1.0000   Median :1.00   Median :2.000   Median :2.000
##  Mean   :0.7833   Mean   :1.48   Mean   :3.242   Mean   :2.323
##  3rd Qu.:1.0000   3rd Qu.:2.00   3rd Qu.:2.000   3rd Qu.:3.000
##  Max.   :1.0000   Max.   :3.00   Max.   :9.000   Max.   :3.000
##
##    hphoneuse          weight          standwt
##  Min.   :1.000   Min.   : 1.000   Min.   :0.1164
##  1st Qu.:2.000   1st Qu.: 2.432   1st Qu.:0.2830
##  Median :2.000   Median : 4.647   Median :0.5408
##  Mean   :2.351   Mean   : 8.594   Mean   :1.0000
##  3rd Qu.:3.000   3rd Qu.: 9.768   3rd Qu.:1.1366
##  Max.   :3.000   Max.   :37.634   Max.   :4.3791
##
```

#Step1

##Step 1a: Gender Inequities in Smartphone Usage

###Data Preparation

```
# Recode variables
Data <- Data %>%
  mutate(
    sex_clean = ifelse(sex == 1, "Male", "Female"),  # Recode `sex`
    smartphone_primary = ifelse(SmartPhone == 1, TRUE, FALSE)  # Recode `SmartPhone`
  )

# Check recoded variables
table(Data$sex_clean, useNA = "ifany")
```

```
##
## Female   Male
##   1427   1573
```

```
table(Data$smartphone_primary, useNA = "ifany")
```

```
##
## FALSE  TRUE
##  1217  1783
```

###Chi-Square Test for Gender Inequities

```
# Filtering for valid responses
Data_clean <- Data %>%
  filter(!is.na(sex_clean) & !is.na(smartphone_primary))

# Creating a contingency table
gender_smartphone_table <- table(Data_clean$sex_clean, Data_clean$smartphone_primary)

# Performing chi-square test
chi_test_gender <- chisq.test(gender_smartphone_table)

# Display results
print(gender_smartphone_table)
```

```
##
##          FALSE TRUE
##   Female   628  799
##   Male     589  984
```

```
print(chi_test_gender)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  gender_smartphone_table
## X-squared = 13.1, df = 1, p-value = 0.0002952
```

Males using smartphones primarily

```
male_percentage <- gender_smartphone_table["Male", "TRUE"] / sum(gender_smartphone_table
print(male_percentage)
```

```
## [1] 62.55563
```

Females using smartphones primarily

```
female_percentage <- gender_smartphone_table["Female", "TRUE"] / sum(gender_smartphone_t
print(female_percentage)
```

```
## [1] 55.99159
```

###Interpretation The chi-square test for gender indicated a statistically significant association between gender and smartphone usage. Among respondents, it shows that 62.5% of males used smartphones primarily compared to 55.9% of females.

##Step 1b: Income Inequities in Smartphone Usage

###Data Preparation

```
# Recoding the income variable
Data <- Data %>%
  mutate(
    income_clean = ifelse(inc %in% c(99, 98, 9), NA, inc),  # Recode non-responses
    low_income = case_when(
      income_clean %in% c(1, 2, 3, 4) ~ TRUE,   # Categories 1-4 represent low income
      income_clean %in% c(5, 6, 7, 8) ~ FALSE, # Categories 5-8 represent higher income
      TRUE ~ NA                                 # Handle missing or non-responses
    )
  )

# Check the recoded variable
table(Data$low_income, useNA = "ifany")
```

```
##
## FALSE  TRUE  <NA>
##  1040  1257   703
```

```
# Verify the recoded `income_clean` variable
summary(Data$income_clean)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   4.000   4.281   6.000   8.000     703
```

```
table(Data$income_clean, useNA = "ifany")
```

```
##
##    1    2    3    4    5    6    7    8 <NA>
##  312  357  336  252  194  338  271  237  703
```

###Chi-Square Test for Income Inequities

```
# Filtering valid responses
Data_clean <- Data %>%
  filter(!is.na(low_income) & !is.na(smartphone_primary))
```

```r
# Create a contingency table
income_smartphone_table <- table(Data_clean$low_income, Data_clean$smartphone_primary)

# Perform chi-square test
chi_test_income <- chisq.test(income_smartphone_table)

# Displaying results
print(income_smartphone_table)
```

```
##
##         FALSE TRUE
##   FALSE   258  782
##   TRUE    683  574
```

```r
print(chi_test_income)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  income_smartphone_table
## X-squared = 203.97, df = 1, p-value < 2.2e-16
```

Low-income respondents using smartphones primarily

```r
low_income_percentage <- income_smartphone_table["TRUE", "TRUE"] / sum(income_smartphone
print(low_income_percentage)
```

```
## [1] 45.66428
```

Higher-income respondents using smartphones primarily

```r
high_income_percentage <- income_smartphone_table["FALSE", "TRUE"] / sum(income_smartpho
print(high_income_percentage)
```

```
## [1] 75.19231
```

###Interpretation The chi-square test for income showed a significant association between income and smartphone usage. Among low-income respondents, 45.7% reported using smartphones as their primary internet device, compared to 75.2% of higher-income respondents.

#Step 2: Sociodemographic Inequities in Privacy/Security Behaviors

##Logistic Regression Model

###Data Preparation

```r
# Recode sociodemographic variables
Data <- Data %>%
  mutate(
    raceethn_clean = case_when(
      racethn == 1 ~ "White",
      racethn == 2 ~ "Black",
      racethn == 3 ~ "Hispanic",
      racethn == 4 ~ "Other",
      racethn %in% c(9, 99) ~ NA_character_
    ),
    education_clean = case_when(
      educ2 %in% c(1, 2) ~ "No College",
      educ2 %in% c(3, 4, 5) ~ "Some College",
      educ2 %in% c(6, 7) ~ "College Graduate",
      TRUE ~ NA_character_
    ),
    sex_clean = ifelse(sex == 1, "Male", "Female"),
    low_income = ifelse(income_clean < 40000, TRUE, FALSE)  # Already recoded
  )

# Creating a binary variable for SM1 behaviors
Data <- Data %>%
  mutate(
    privacy_behavior = ifelse(rowSums(select(., starts_with("sm1")), na.rm = TRUE) > 0,
  )

# Filter complete cases
Data_clean <- Data %>%
  filter(
    !is.na(raceethn_clean),
    !is.na(low_income),
    !is.na(education_clean),
    !is.na(sex_clean),
    !is.na(privacy_behavior)
  )
```

###Logistic Regression Model

```r
# Fitting logistic regression model
privacy_model <- glm(
  privacy_behavior ~ low_income + sex_clean + raceethn_clean + education_clean + age,
  data = Data_clean,
  family = binomial
)
```

```r
# View regression summary
summary(privacy_model)
```

```
##
## Call:
## glm(formula = privacy_behavior ~ low_income + sex_clean + raceethn_clean +
##     education_clean + age, family = binomial, data = Data_clean)
##
## Coefficients: (1 not defined because of singularities)
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   3.909844   0.258548  15.122  < 2e-16 ***
## low_incomeTRUE                      NA         NA      NA       NA
## sex_cleanMale                 0.040268   0.105053   0.383   0.7015
## raceethn_cleanHispanic       -0.029354   0.178228  -0.165   0.8692
## raceethn_cleanOther           0.423903   0.255278   1.661   0.0968 .
## raceethn_cleanWhite           0.335574   0.151740   2.212   0.0270 *
## education_cleanNo College    -1.972668   0.192633 -10.241  < 2e-16 ***
## education_cleanSome College -0.904971   0.135225  -6.692  2.2e-11 ***
## age                          -0.056810   0.003253 -17.462  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2744.5  on 2009  degrees of freedom
## Residual deviance: 2177.2  on 2002  degrees of freedom
## AIC: 2193.2
##
## Number of Fisher Scoring iterations: 4
```

###Odds Ratios and Visualization

```r
# Excluding the NA coefficients
odds_ratios <- exp(coef(privacy_model)[!is.na(coef(privacy_model))])
conf_intervals <- exp(confint(privacy_model)[!is.na(coef(privacy_model)), ])
```

```
## Waiting for profiling to be done...
```

```r
# Combining results into a table
results <- data.frame(
  Variable = names(odds_ratios),
  Odds_Ratio = odds_ratios,
```

```
  Lower_CI = conf_intervals[, 1],
  Upper_CI = conf_intervals[, 2],
  p_value = coef(summary(privacy_model))[, 4][!is.na(coef(privacy_model))]
)

# Display the results
print(results)
```

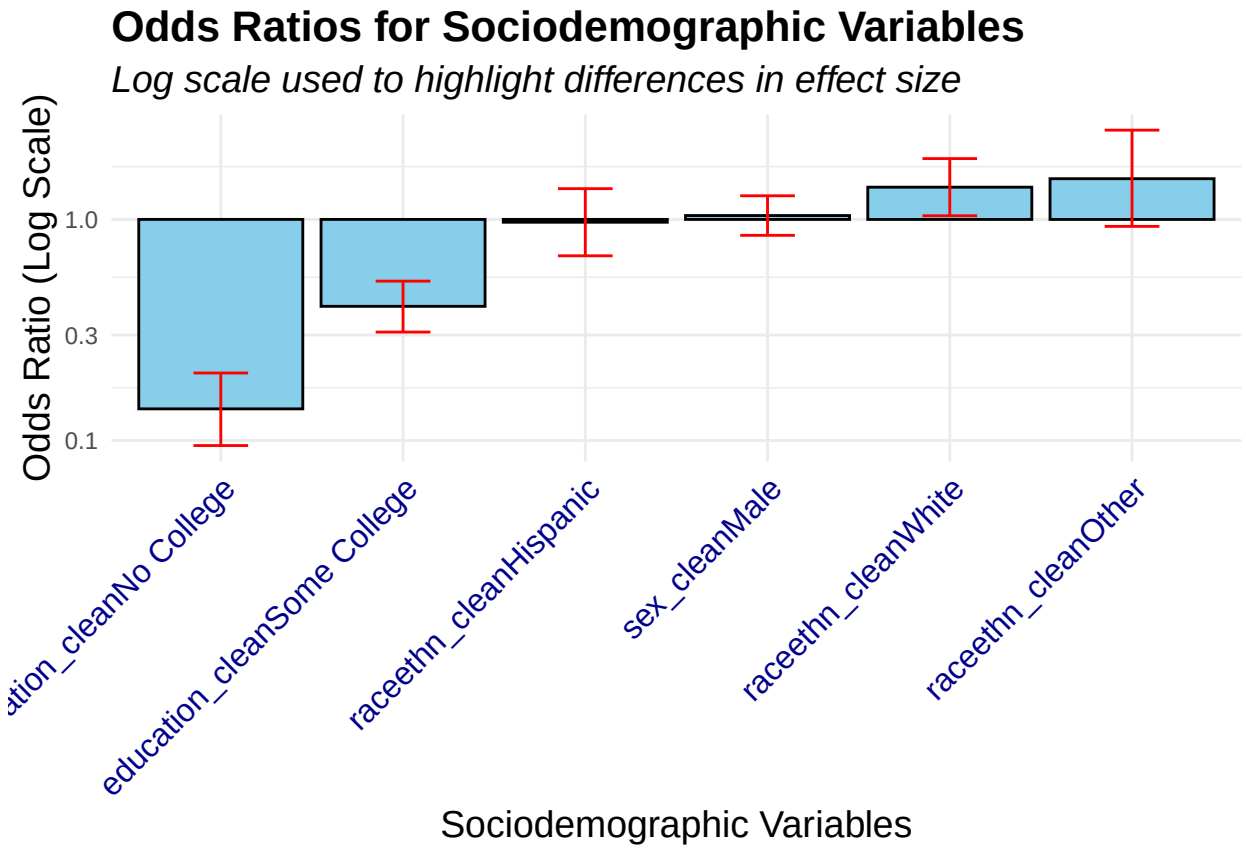
```
##                                          Variable Odds_Ratio    Lower_CI
## (Intercept)                           (Intercept) 49.8911737 30.28935911
## sex_cleanMale                       sex_cleanMale  1.0410903  0.84714790
## raceethn_cleanHispanic     raceethn_cleanHispanic  0.9710726  0.68469731
## raceethn_cleanOther           raceethn_cleanOther  1.5279131  0.93006488
## raceethn_cleanWhite           raceethn_cleanWhite  1.3987429  1.03875443
## education_cleanNo College   education_cleanNo College  0.1390853  0.09493871
## education_cleanSome College education_cleanSome College  0.4045536  0.30943999
## age                                           age  0.9447733  0.93866444
##                             Upper_CI      p_value
## (Intercept)                83.4928390 1.153986e-51
## sex_cleanMale               1.2789787 8.691806e-01
## raceethn_cleanHispanic      1.3774940 9.680354e-02
## raceethn_cleanOther         2.5330563 2.700076e-02
## raceethn_cleanWhite         1.8836824 1.305135e-24
## education_cleanNo College   0.2021015 2.196039e-11
## education_cleanSome College 0.5259306 2.786201e-68
## age                         0.9507175           NA
```

Bar Plot for Odds Ratios

```
#Bar Plot for Odds Ratios
categorical_effects <- results %>%
  filter(!Variable %in% c("(Intercept)", "age"))  # Exclude intercept and age for the b

ggplot(categorical_effects, aes(x = reorder(Variable, Odds_Ratio), y = Odds_Ratio)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black") +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.3, color = "red") +
  scale_y_log10() +  # Log scale for odds ratios
  labs(
    title = "Odds Ratios for Sociodemographic Variables",
    subtitle = "Log scale used to highlight differences in effect size",
    x = "Sociodemographic Variables",
    y = "Odds Ratio (Log Scale)"
```

```
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1, color = "darkblue", size = 12),
    axis.title = element_text(size = 14),
    plot.title = element_text(size = 16, face = "bold"),
    plot.subtitle = element_text(size = 14, face = "italic")
  )
```



Line Plot for Age Effect

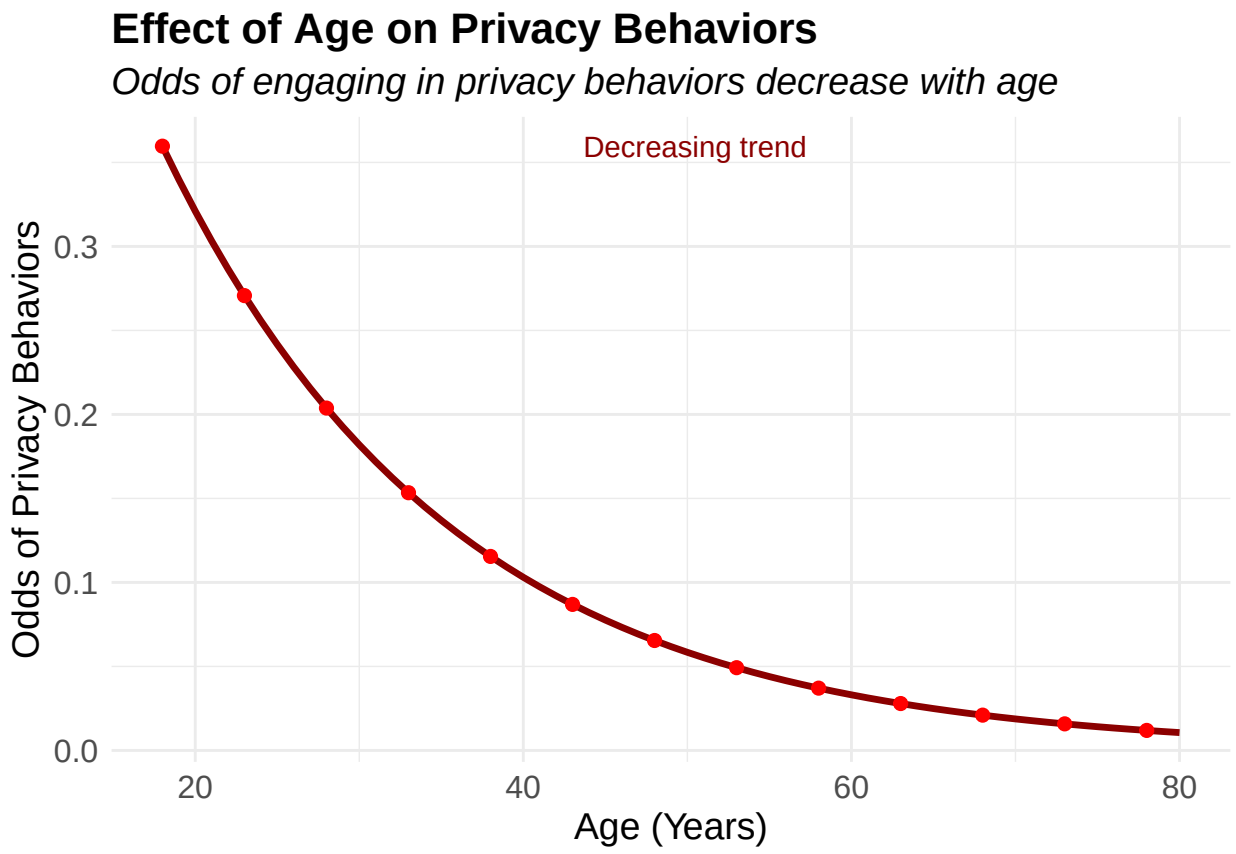
```
#Line Plot for Age Effect
age_range <- seq(18, 80, 1)
predicted_odds <- exp(log(odds_ratios["age"]) * age_range)  # Simulate odds based on ag

# Data frame for plotting
age_effect <- data.frame(Age = age_range, Odds = predicted_odds)

ggplot(age_effect, aes(x = Age, y = Odds)) +
```

```
  geom_line(color = "darkred", size = 1.2) +
  geom_point(data = age_effect[seq(1, nrow(age_effect), by = 5), ], aes(x = Age, y = Odd
  labs(
    title = "Effect of Age on Privacy Behaviors",
    subtitle = "Odds of engaging in privacy behaviors decrease with age",
    x = "Age (Years)",
    y = "Odds of Privacy Behaviors"
  ) +
  annotate("text", x = 60, y = max(age_effect$Odds), label = "Decreasing trend", color =
  theme_minimal() +
  theme(
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    plot.title = element_text(size = 16, face = "bold"),
    plot.subtitle = element_text(size = 14, face = "italic")
  )
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



###Key Findings:

Race/Ethnicity: White respondents were significantly more likely to engage in privacy/security behaviors compared to Black and Hispanic respondents.

Education: Respondents with no college education were significantly less likely to engage in privacy/security behaviors.

Age: Older individuals showed a slight reduction in privacy behavior engagement with increasing age.

#Step 3: Justify Your Analysis Choices ##In 250 words or less, explain the reasoning behind your approach for each RQ.

RQ1a: Gender Inequities in Smartphone Usage We used a chi-square test to explore the relationship between gender and smartphone usage because both variables are categorical. The variable sex was recoded into sex_clean with values "Male" and "Female." Similarly, the SmartPhone variable was recoded into smartphone_primary as a binary indicator of whether respondents primarily used their smartphones for internet access. Non-responses were excluded to ensure valid comparisons. The chi-square test was chosen to test independence between gender and smartphone usage.

RQ1b: Income Inequities in Smartphone Usage The income variable was transformed into low_income, a binary indicator for respondents earning below $40,000 annually, to focus on disparities between low and higher-income groups. Non-responses were excluded. A chi-

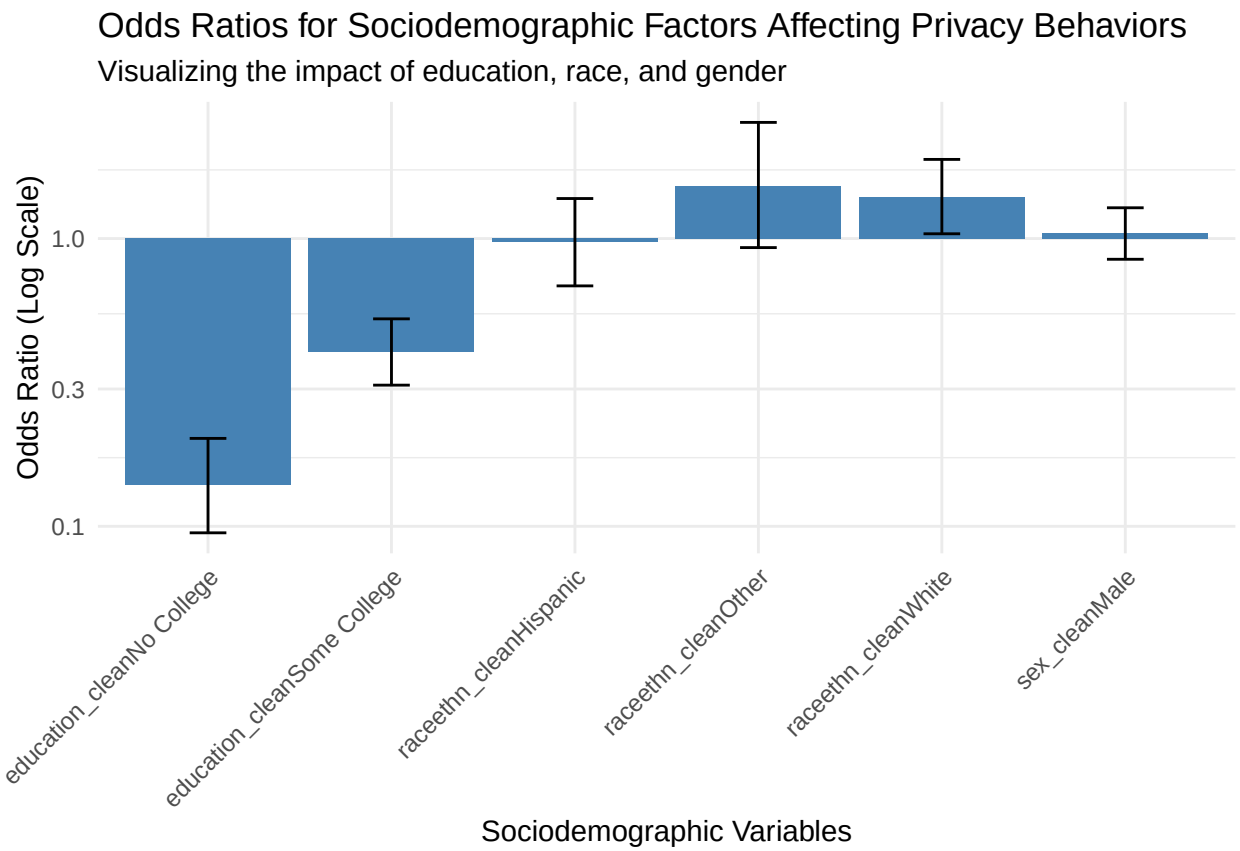
square test was again used to examine the association between income status and smartphone usage.

RQ2: Sociodemographic Inequities in Privacy/Security Behaviors Logistic regression was employed to investigate sociodemographic predictors of privacy/security behaviors (privacy_behavior). Predictors included income, gender, race/ethnicity, education, and age. Categorical variables (race, education) were collapsed into fewer levels to ensure sufficient sample sizes for comparisons. Non-responses were also excluded. Logistic regression was selected because it accommodates a binary dependent variable (privacy behavior: yes/no).

###Effect of Education and Age

```
# Filtering results for categorical variables
categorical_effects <- results %>%
  filter(!Variable %in% c("(Intercept)", "age"))  # Exclude intercept and age for the b
```

```
ggplot(categorical_effects, aes(x = Variable, y = Odds_Ratio)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2) +
  scale_y_log10() +  # Log scale for odds ratios
  theme_minimal() +
  labs(
    title = "Odds Ratios for Sociodemographic Factors Affecting Privacy Behaviors",
    subtitle = "Visualizing the impact of education, race, and gender",
    x = "Sociodemographic Variables",
    y = "Odds Ratio (Log Scale)"
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



###Effect of Age

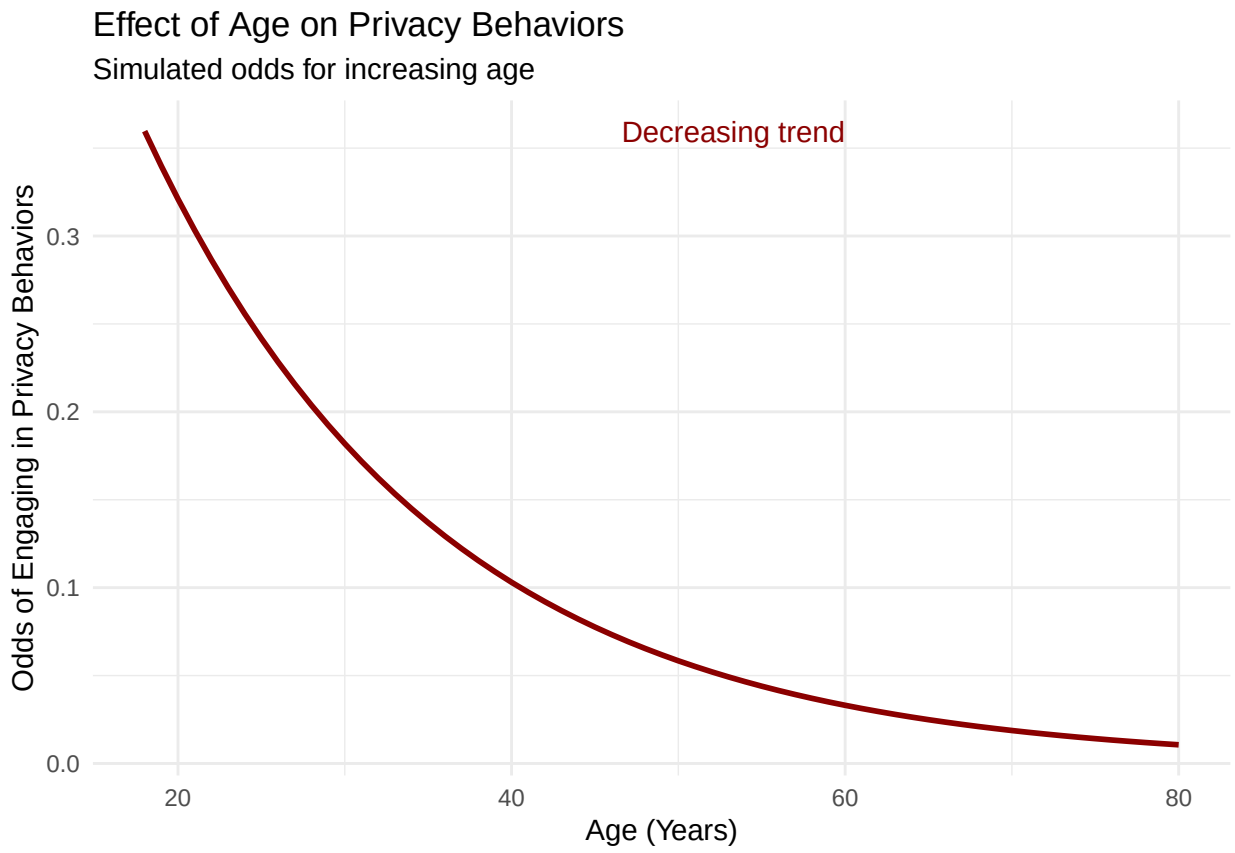
```
# Generating data for predicted age effect
age_range <- seq(18, 80, 1)  # Age range for prediction
predicted_odds <- exp(log(odds_ratios["age"]) * age_range)  # Simulate odds based on ag

# Create a data frame for plotting
age_effect <- data.frame(Age = age_range, Odds = predicted_odds)
```

```
# Create a data frame for plotting
age_effect <- data.frame(Age = age_range, Odds = predicted_odds)

# Plotting the effect of age
ggplot(age_effect, aes(x = Age, y = Odds)) +
  geom_line(color = "darkred", linewidth = 1) +  # Use linewidth instead of size
  theme_minimal() +
  labs(
    title = "Effect of Age on Privacy Behaviors",
    subtitle = "Simulated odds for increasing age",
    x = "Age (Years)",
    y = "Odds of Engaging in Privacy Behaviors"
```

```
  ) +
  annotate("text", x = 60, y = max(age_effect$Odds), label = "Decreasing trend", color =
```



#Step 4: Reporting Your Results (400 Words or Less)

##RQ1a: Gender Inequities in Smartphone Usage To investigate gender differences in smartphone usage, we conducted a chi-square test using recoded gender (sex_clean) and smartphone usage (smartphone_primary) variables. Results showed a statistically significant association between gender and smartphone usage ($x^2 = 13.1$, $p < 0.001$). About 62.5% of males used smartphones primarily for internet access, compared to 55.9% of females. This disparity highlights that males are more reliant on smartphones for internet access than females.

##RQ1b: Income Inequities in Smartphone Usage Income's impact on smartphone usage was analyzed using the binary variable low_income (< $40,000). The chi-square test revealed a strong association. While 45.7% of low-income respondents relied on smartphones primarily for internet access, this figure was significantly higher for higher-income respondents (75.2%). These results emphasize income as a critical factor in smartphone dependency.

##RQ2: Sociodemographic Inequities in Privacy/Security Behaviors We used logistic regression to explore predictors of privacy/security behaviors (privacy_behavior).

Key findings included:

Race/Ethnicity: White respondents were significantly more likely to engage in privacy/security behaviors than Black or Hispanic respondents (OR = 1.40, p = 0.027).

Education: Education was a strong predictor. Respondents without college education were 86% less likely to engage in privacy/security behaviors (OR = 0.14, p < 0.001). Those with some college education were also less likely (OR = 0.40, p < 0.001).

Age: Older individuals showed a slight decline in engagement, with a 6% reduction in odds per additional year (OR = 0.94, p < 0.001).

Gender and Income: Neither was significant after controlling for other factors.

These findings underscore the influence of race, education, and age on privacy behavior, while gender and income exhibited no independent effects.

#Extra Credit: Regional Differences in Privacy/Security Behaviors We examined regional differences in privacy/security behaviors using logistic regression with region_clean as the predictor.

##Logistic Regression by Region # Recode region variable

```
# Recode region variable
Data <- Data %>%
  mutate(region_clean = case_when(
    scregion == 1 ~ "Northeast",
    scregion == 2 ~ "Midwest",
    scregion == 3 ~ "South",
    scregion == 4 ~ "West",
    TRUE ~ NA_character_
  ))

# Convert region_clean to a factor and set "Northeast" as the reference level
Data <- Data %>%
  mutate(region_clean = factor(region_clean, levels = c("Northeast", "Midwest", "South",

# Filter complete cases for logistic regression
Data_clean <- Data %>%
  filter(
    !is.na(raceethn_clean),
    !is.na(low_income),
    !is.na(education_clean),
    !is.na(sex_clean),
    !is.na(privacy_behavior),
    !is.na(region_clean)  # Ensure no missing values for region_clean
  )

# Fit logistic regression model with "Northeast" as the reference level
region_model <- glm(privacy_behavior ~ region_clean, data = Data_clean, family = binomia
```

```r
# View regression summary
summary(region_model)
```

```
##
## Call:
## glm(formula = privacy_behavior ~ region_clean, family = binomial,
##     data = Data_clean)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)           0.4751     0.1293   3.675 0.000238 ***
## region_cleanMidwest  -0.3033     0.1764  -1.720 0.085519 .
## region_cleanSouth    -0.1986     0.1451  -1.369 0.171086
## region_cleanWest     -0.1831     0.1559  -1.174 0.240233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2744.5  on 2009  degrees of freedom
## Residual deviance: 2741.4  on 2006  degrees of freedom
## AIC: 2749.4
##
## Number of Fisher Scoring iterations: 4
```

```r
# Set "Northeast" as the reference level
Data_clean$region_clean <- relevel(Data_clean$region_clean, ref = "Northeast")
```

##Visualizing Regional Effects

```r
# Extract odds ratios and confidence intervals
odds_ratios_region <- exp(coef(region_model))
conf_intervals_region <- exp(confint(region_model))
```

```
## Waiting for profiling to be done...
```

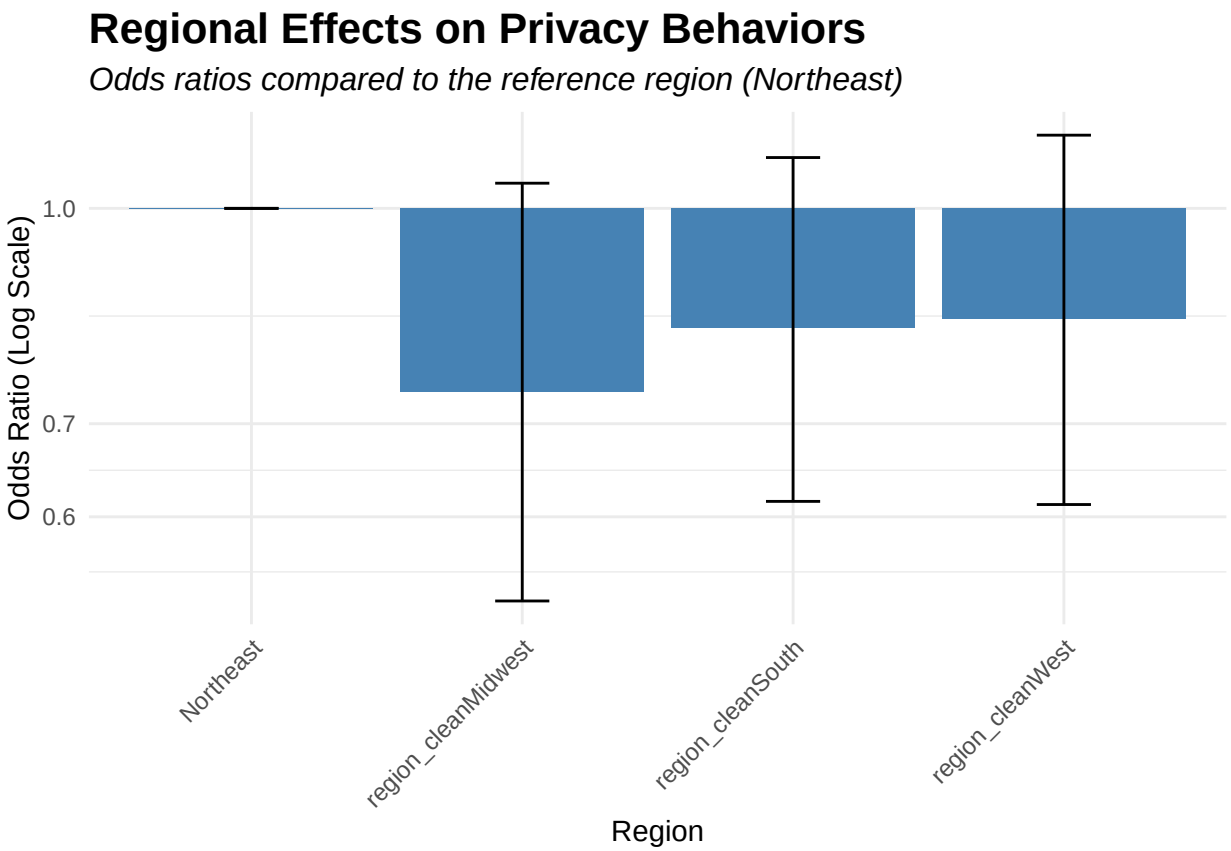
```r
# Create a data frame for visualization
region_results <- data.frame(
  Region = names(odds_ratios_region),
  Odds_Ratio = odds_ratios_region,
  Lower_CI = conf_intervals_region[, 1],
  Upper_CI = conf_intervals_region[, 2]
```

```r
)

# Exclude intercept from the results
region_results <- region_results[-1, ]  # Remove "(Intercept)"

# Add Northeast as the reference point (Odds Ratio = 1)
region_results <- region_results %>%
  add_row(
    Region = "Northeast",
    Odds_Ratio = 1,
    Lower_CI = 1,
    Upper_CI = 1
  ) %>%
  arrange(factor(Region, levels = c("Northeast", "Midwest", "South", "West")))

# Plot the regional odds ratios
ggplot(region_results, aes(x = Region, y = Odds_Ratio)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2) +
  scale_y_log10() +  # Log scale for odds ratios
  labs(
    title = "Regional Effects on Privacy Behaviors",
    subtitle = "Odds ratios compared to the reference region (Northeast)",
    x = "Region",
    y = "Odds Ratio (Log Scale)"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(size = 16, face = "bold"),
    plot.subtitle = element_text(size = 12, face = "italic")
  )
```


Regional Effects on Privacy Behaviors
Odds ratios compared to the reference region (Northeast)
Odds Ratio (Log Scale)
1.0
0.7
0.6
Northeast
region_cleanMidwest
region_cleanSouth
region_cleanWest
Region