Working report

Contents

Survey questions

Q1. Before receiving this survey, did you know influenza is different from the stomach flu?

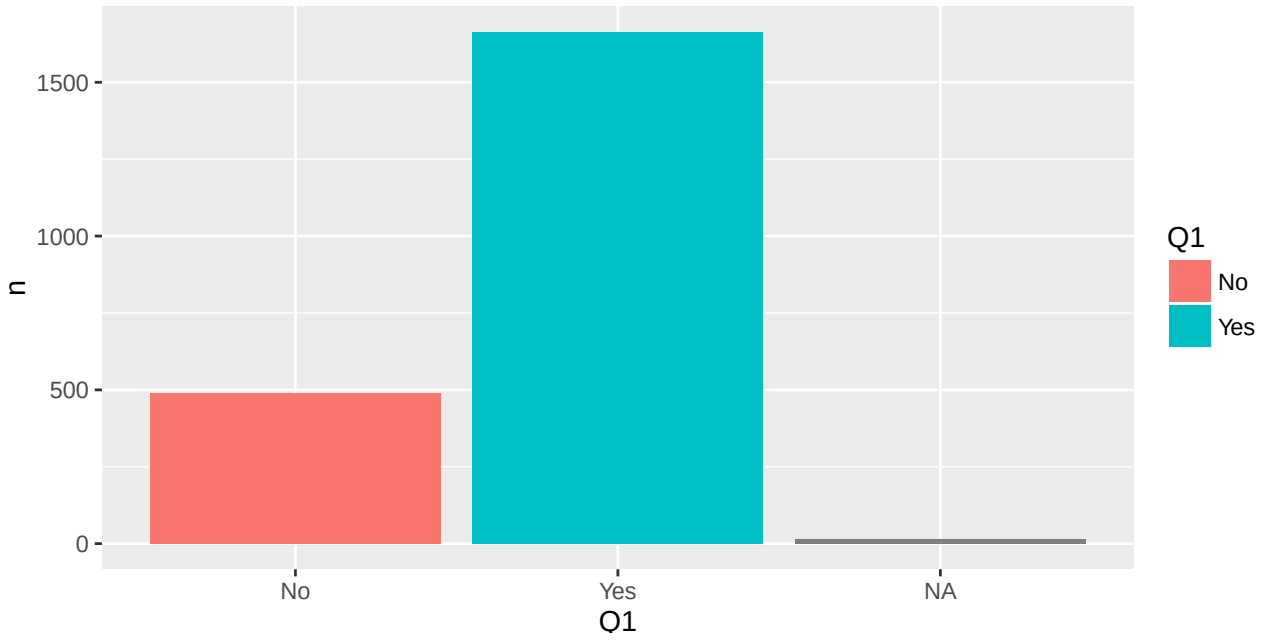
```
# Q1 summary
with(data, table(Q1))
```

```
## Q1
##   No  Yes
##  488 1664
```

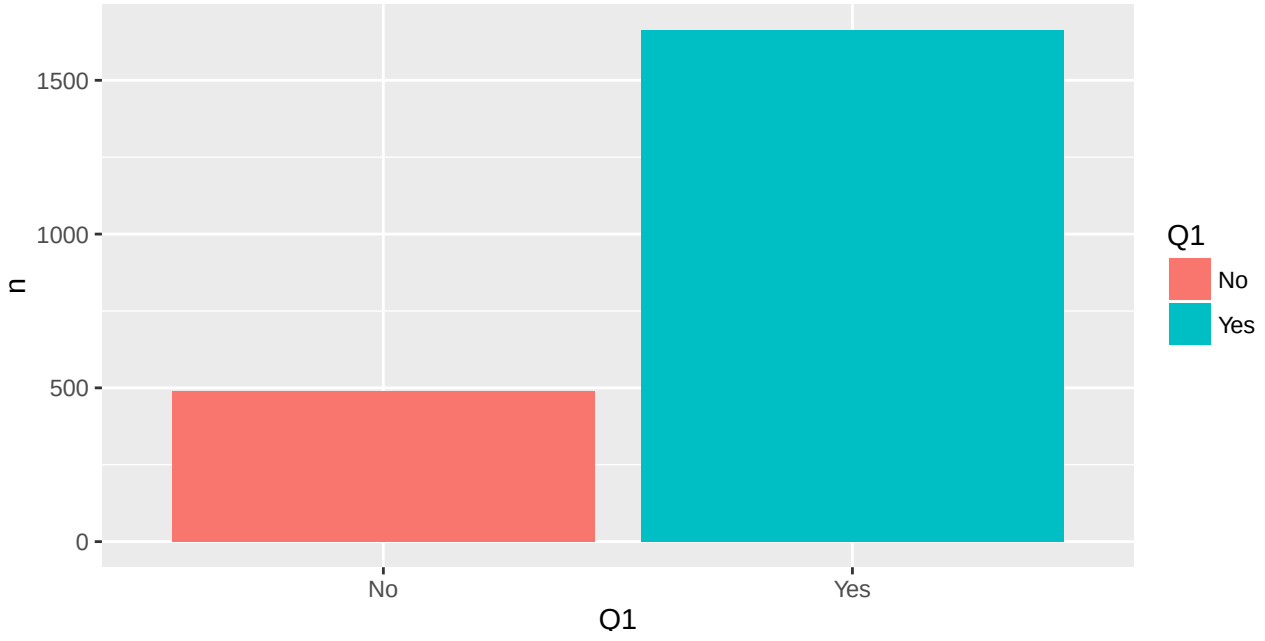
```
(
q1 <- data %>%
  count(Q1)
)
```

```
## Source: local data frame [3 x 2]
##
##       Q1     n
##   <fctr> <int>
## 1     No   488
## 2    Yes  1664
## 3     NA    16
```

```
# plot
ggplot(q1, aes(x = Q1, y = n, fill = Q1)) + geom_bar(stat = 'identity')
```




```
# plot without na's
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = Q1)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



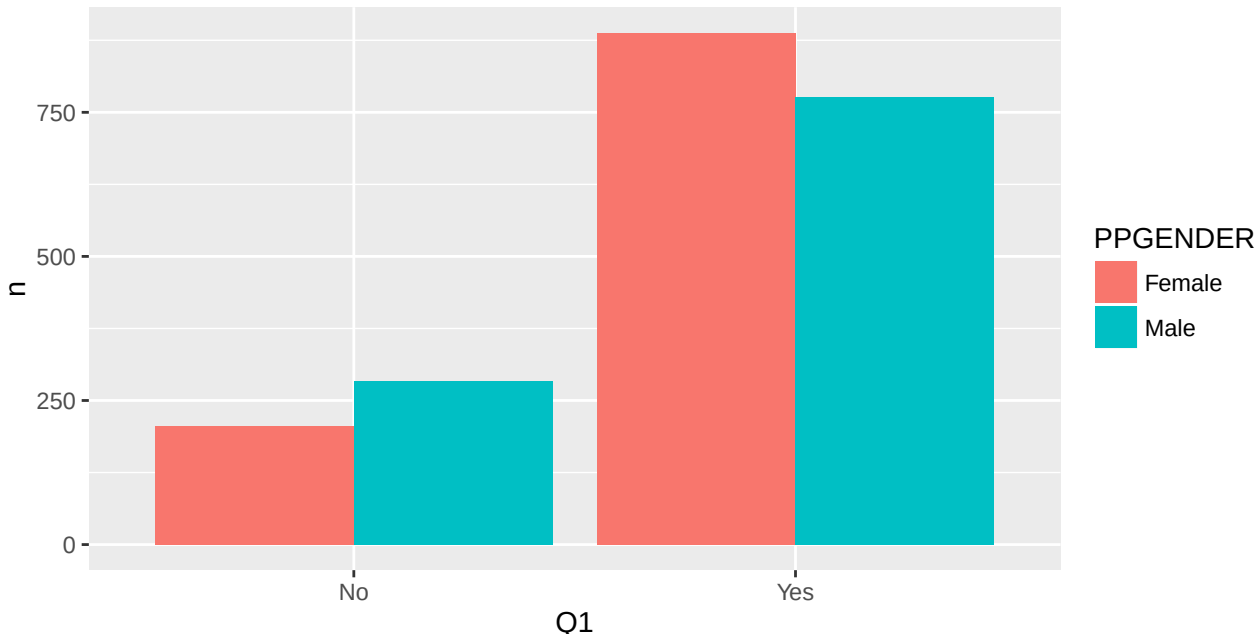

```
# by gender, PPGENDER
with(data, table(PPGENDER, Q1))
```

```
##          Q1
## PPGENDER  No Yes
##   Female 205 888
##   Male   283 776
```

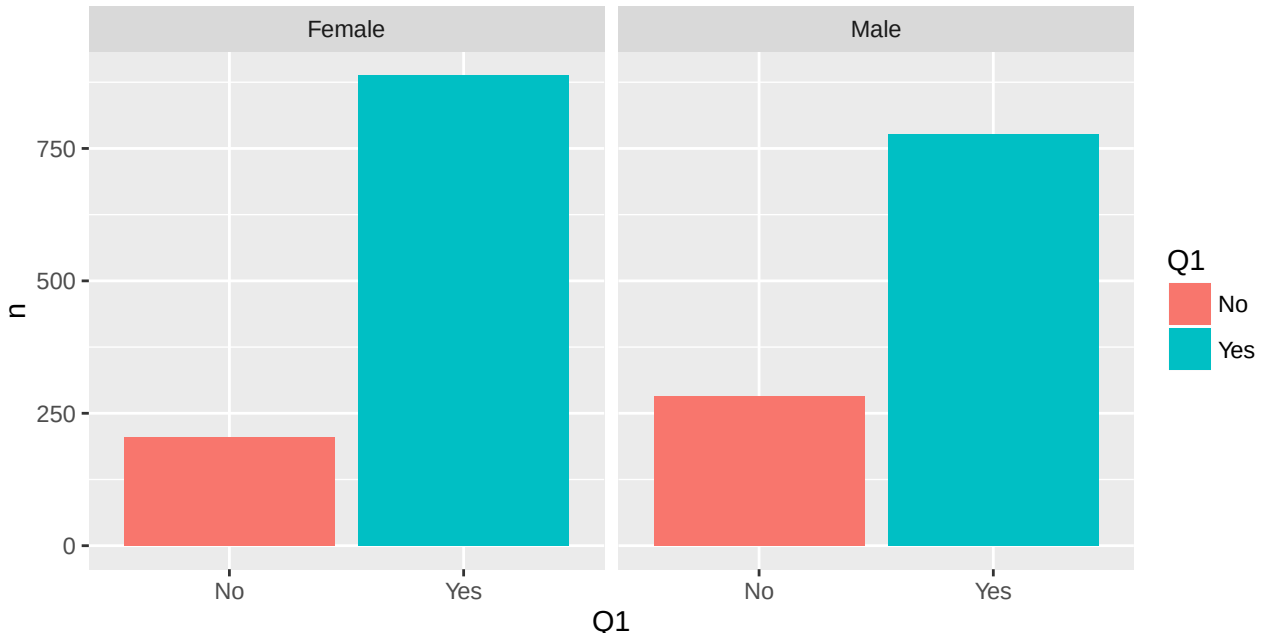
```
(
q1 <- data %>%
  count(Q1, PPGENDER)
)
```

```
## Source: local data frame [6 x 3]
## Groups: Q1 [?]
##
##       Q1 PPGENDER     n
##   (fctr)   (fctr) (int)
## 1     No   Female   205
## 2     No     Male   283
## 3    Yes   Female   888
## 4    Yes     Male   776
## 5     NA   Female     4
## 6     NA     Male    12
```

```
# plot
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# plot with facet
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = Q1)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~PPGENDER)
```




```
# by ethnicity, PPETHM
with(data, table(PPETHM, Q1))
```

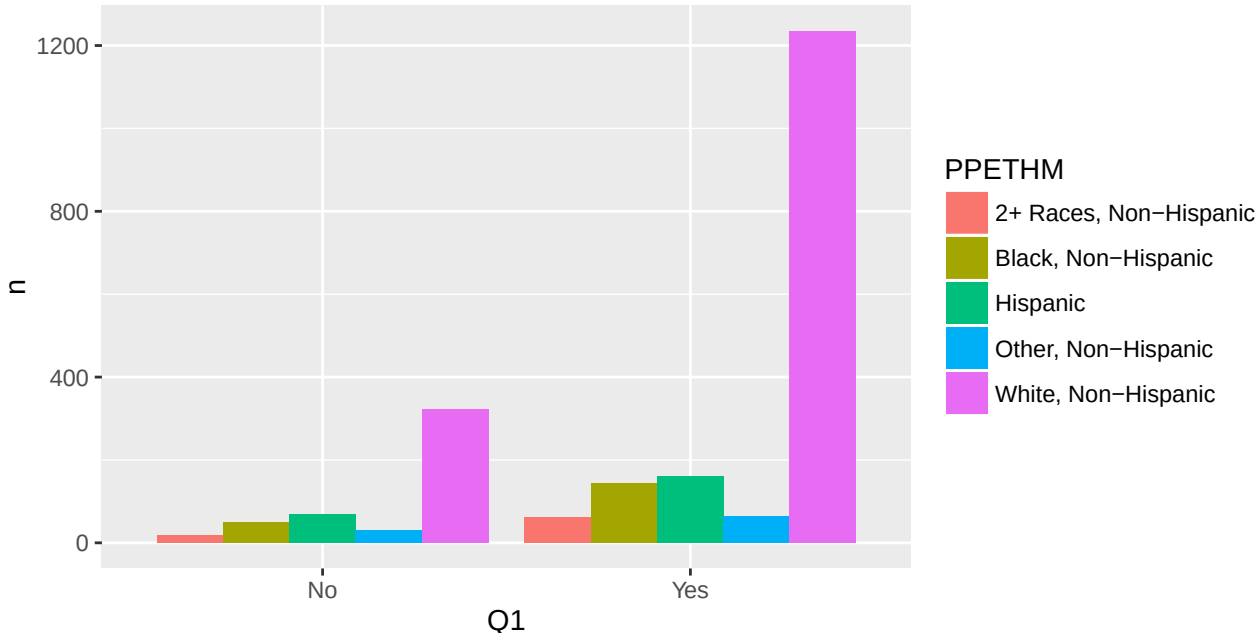
```
##                          Q1
## PPETHM                     No  Yes
##   2+ Races, Non-Hispanic   18   62
##   Black, Non-Hispanic      50  143
##   Hispanic                 69  161
##   Other, Non-Hispanic      29   63
##   White, Non-Hispanic     322 1235
```

```
(
q1 <- data %>%
  count(Q1, PPETHM)
)
```

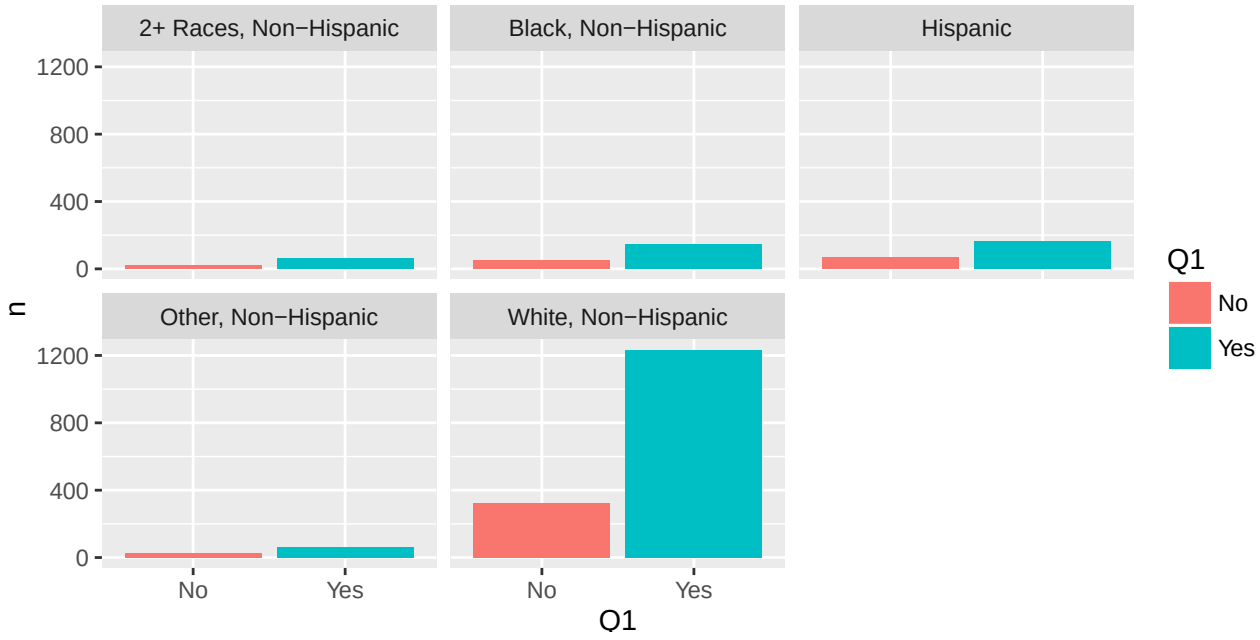
```
## Source: local data frame [14 x 3]
## Groups: Q1 [?]
##
##        Q1                 PPETHM     n
##    (fctr)                 (fctr) (int)
## 1      No 2+ Races, Non-Hispanic    18
## 2      No    Black, Non-Hispanic    50
## 3      No               Hispanic    69
## 4      No    Other, Non-Hispanic    29
## 5      No    White, Non-Hispanic   322
## 6     Yes 2+ Races, Non-Hispanic    62
## 7     Yes    Black, Non-Hispanic   143
## 8     Yes               Hispanic   161
## 9     Yes    Other, Non-Hispanic    63
## 10    Yes    White, Non-Hispanic  1235
## 11     NA    Black, Non-Hispanic     2
```

```
## 12     NA                Hispanic     2
## 13     NA    Other, Non-Hispanic     1
## 14     NA    White, Non-Hispanic    11
```

```r
# plot
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```r
# plot with facet
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = Q1)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~PPETHM)
```



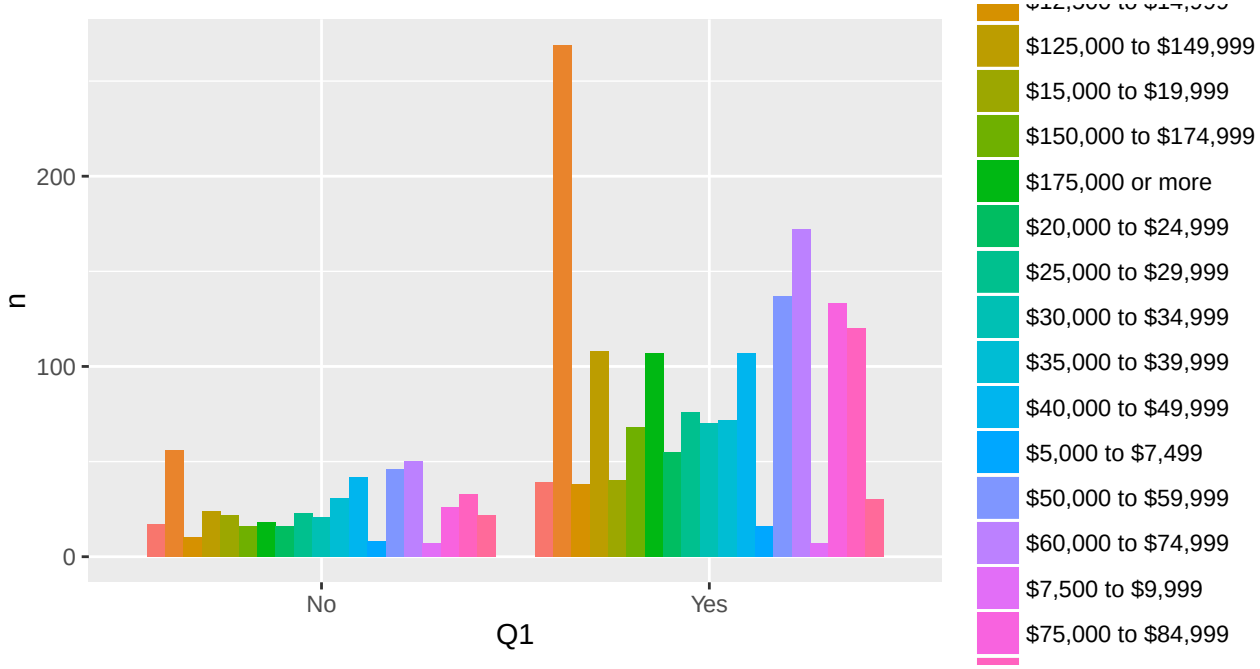
```r
# by income, PPINCIMP
with(data, table(PPINCIMP, Q1))
```

```
##                       Q1
## PPINCIMP                No Yes
##   $10,000 to $12,499    17  39
##   $100,000 to $124,999  56 269
##   $12,500 to $14,999    10  38
##   $125,000 to $149,999  24 108
##   $15,000 to $19,999    22  40
##   $150,000 to $174,999  16  68
##   $175,000 or more      18 107
##   $20,000 to $24,999    16  55
##   $25,000 to $29,999    23  76
##   $30,000 to $34,999    21  70
##   $35,000 to $39,999    31  72
##   $40,000 to $49,999    42 107
##   $5,000 to $7,499       8  16
##   $50,000 to $59,999    46 137
##   $60,000 to $74,999    50 172
##   $7,500 to $9,999       7   7
##   $75,000 to $84,999    26 133
##   $85,000 to $99,999    33 120
##   Less than $5,000      22  30
```

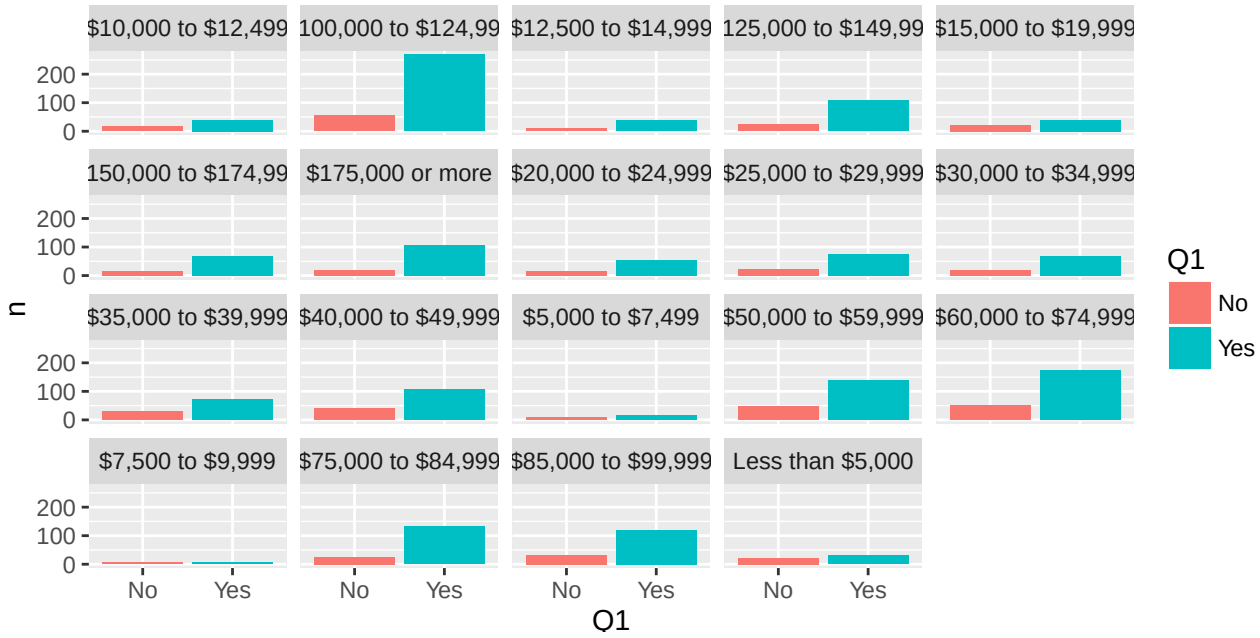
```r
(
q1 <- data %>%
  count(Q1, PPINCIMP)
)
```

```
## Source: local data frame [50 x 3]
## Groups: Q1 [?]
##
##        Q1              PPINCIMP     n
##    (fctr)                (fctr) (int)
## 1      No   $10,000 to $12,499    17
## 2      No $100,000 to $124,999    56
## 3      No   $12,500 to $14,999    10
## 4      No $125,000 to $149,999    24
## 5      No   $15,000 to $19,999    22
## 6      No $150,000 to $174,999    16
## 7      No     $175,000 or more    18
## 8      No   $20,000 to $24,999    16
## 9      No   $25,000 to $29,999    23
## 10     No   $30,000 to $34,999    21
## ..    ...                   ...   ...
```

```r
# plot
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# plot with facet
ggplot(q1[!is.na(q1$Q1), ], aes(x = Q1, y = n, fill = Q1)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~PPINCIMP)
```



Q2. Have you had an illness with influenza-like symptoms since August 2015?

```
#
with(data, table(Q2))
```

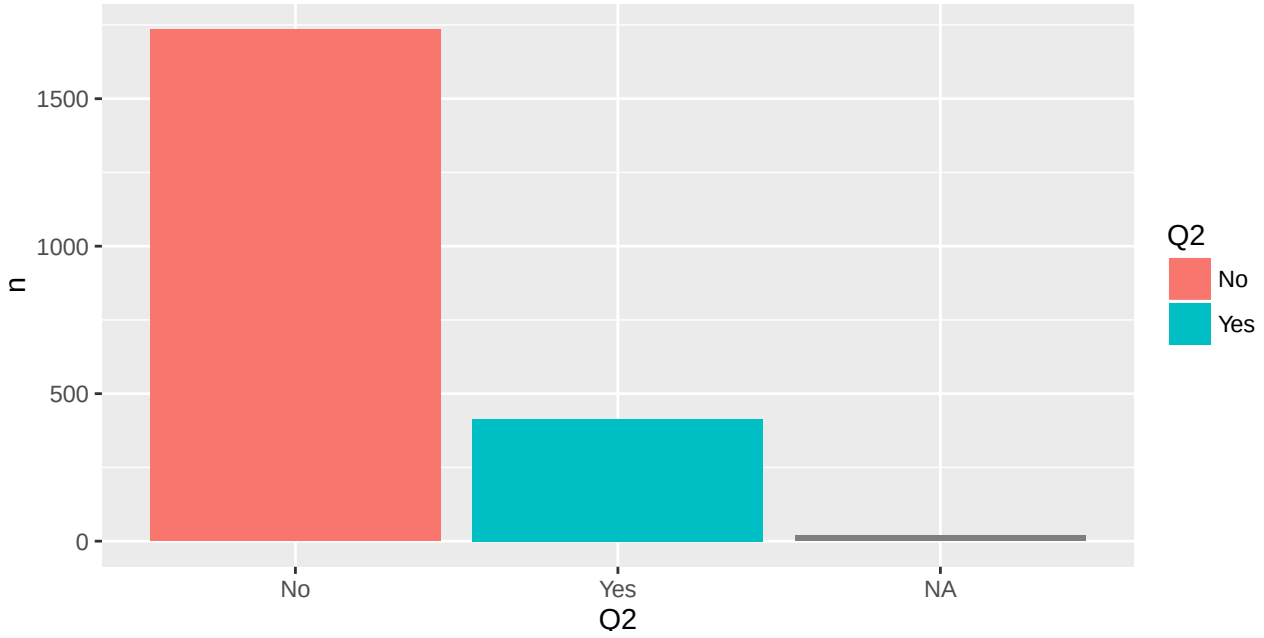
```
## Q2
```

```
##   No  Yes
## 1735  414
```

```
(
q2 <- data %>%
  count(Q2)
)
```

```
## Source: local data frame [3 x 2]
##
##       Q2     n
##   <fctr> <int>
## 1     No  1735
## 2    Yes   414
## 3     NA    19
```

```
ggplot(q2, aes(x = Q2, y = n, fill = Q2)) + geom_bar(stat = 'identity')
```



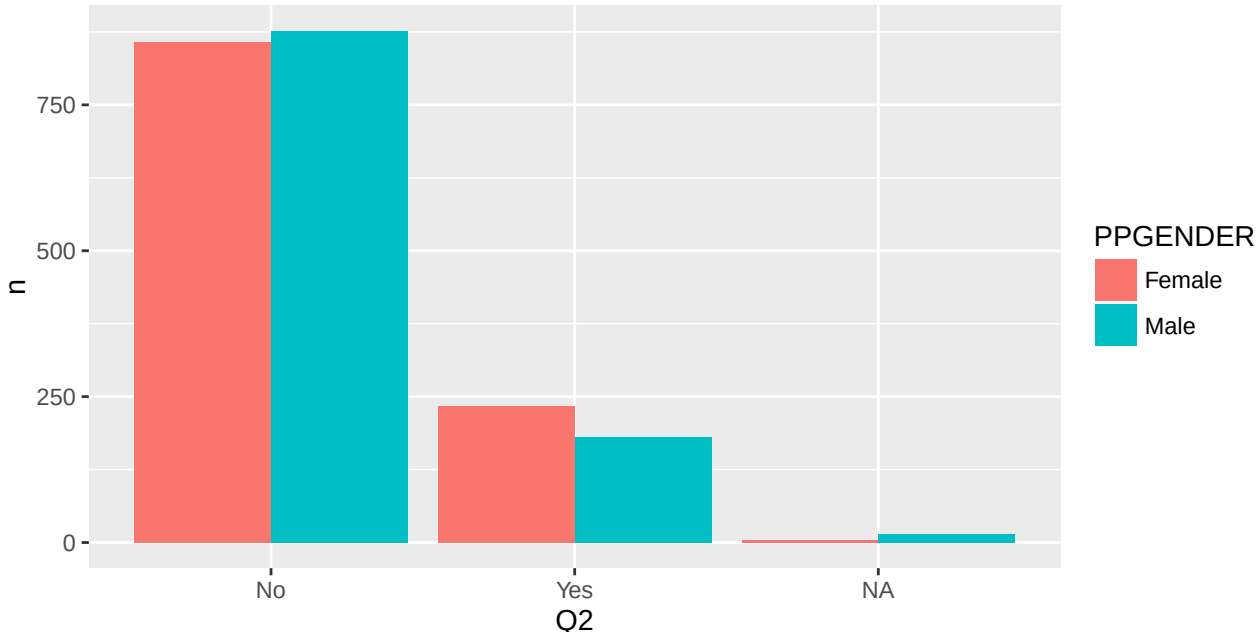

```
# by gender
with(data, table(Q2, PPGENDER))
```

```
##      PPGENDER
## Q2    Female Male
##   No     858  877
##   Yes    234  180
```

```
(
q2 <- data %>%
  count(Q2, PPGENDER)
)
```

```
## Source: local data frame [6 x 3]
## Groups: Q2 [?]
##
##       Q2 PPGENDER     n
##   (fctr)   (fctr) (int)
## 1     No   Female   858
## 2     No     Male   877
## 3    Yes   Female   234
## 4    Yes     Male   180
## 5     NA   Female     5
## 6     NA     Male    14
```

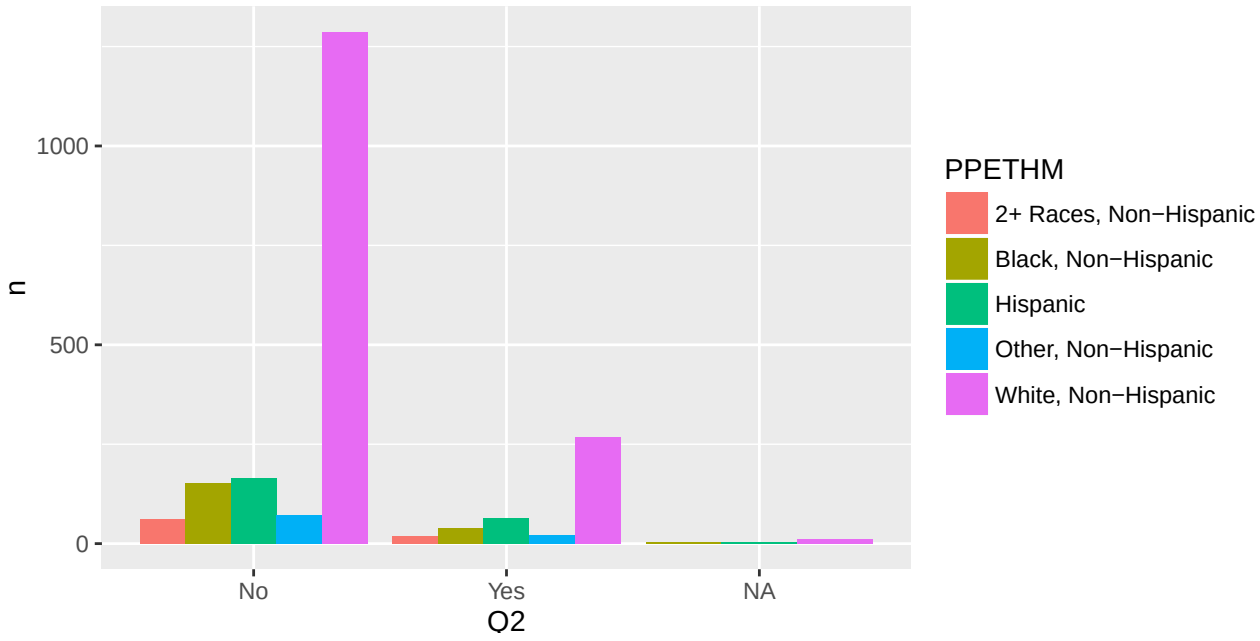
```
ggplot(q2, aes(x = Q2, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# by ethnicity
with(data, table(Q2, PPETHM))
```

```
##      PPETHM
## Q2    2+ Races, Non-Hispanic Black, Non-Hispanic Hispanic
##   No                        61                 152      164
##   Yes                       19                  39       65
##      PPETHM
## Q2    Other, Non-Hispanic White, Non-Hispanic
##   No                    71                1287
##   Yes                   22                 269
```

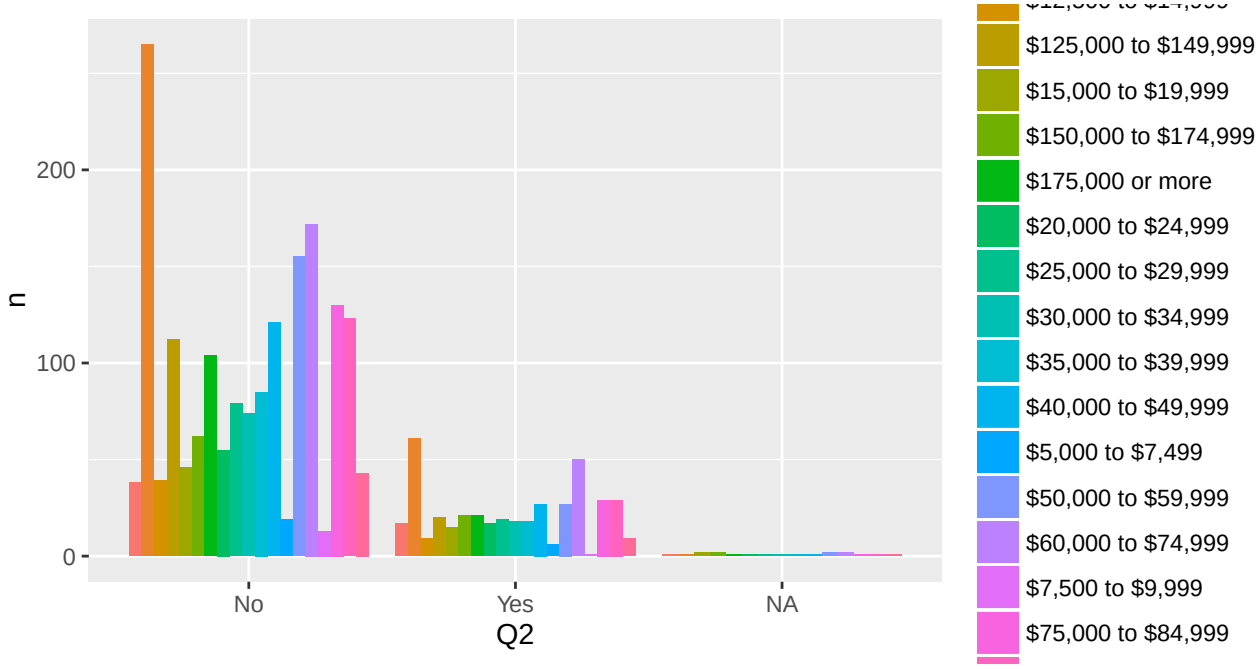
```
q2 <- data %>%
  count(Q2, PPETHM)
ggplot(q2, aes(x = Q2, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# by income
with(data, table(Q2, PPINCIMP))
```

```
##      PPINCIMP
## Q2    $10,000 to $12,499 $100,000 to $124,999 $12,500 to $14,999
##   No                  38                  265                 39
##   Yes                 17                   61                  9
##      PPINCIMP
## Q2    $125,000 to $149,999 $15,000 to $19,999 $150,000 to $174,999
##   No                   112                 46                   62
##   Yes                   20                 15                   21
##      PPINCIMP
## Q2    $175,000 or more $20,000 to $24,999 $25,000 to $29,999
##   No               104                 55                 79
##   Yes               21                 17                 19
##      PPINCIMP
## Q2    $30,000 to $34,999 $35,000 to $39,999 $40,000 to $49,999
##   No                  74                 85                121
##   Yes                 18                 18                 27
##      PPINCIMP
## Q2    $5,000 to $7,499 $50,000 to $59,999 $60,000 to $74,999
##   No                19                155                172
##   Yes                6                 27                 50
##      PPINCIMP
## Q2    $7,500 to $9,999 $75,000 to $84,999 $85,000 to $99,999
##   No                13                130                123
##   Yes                1                 29                 29
##      PPINCIMP
## Q2    Less than $5,000
##   No                43
##   Yes                9
```

```
q2 <- data %>%
  count(Q2, PPINCIMP)
ggplot(q2, aes(x = Q2, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```

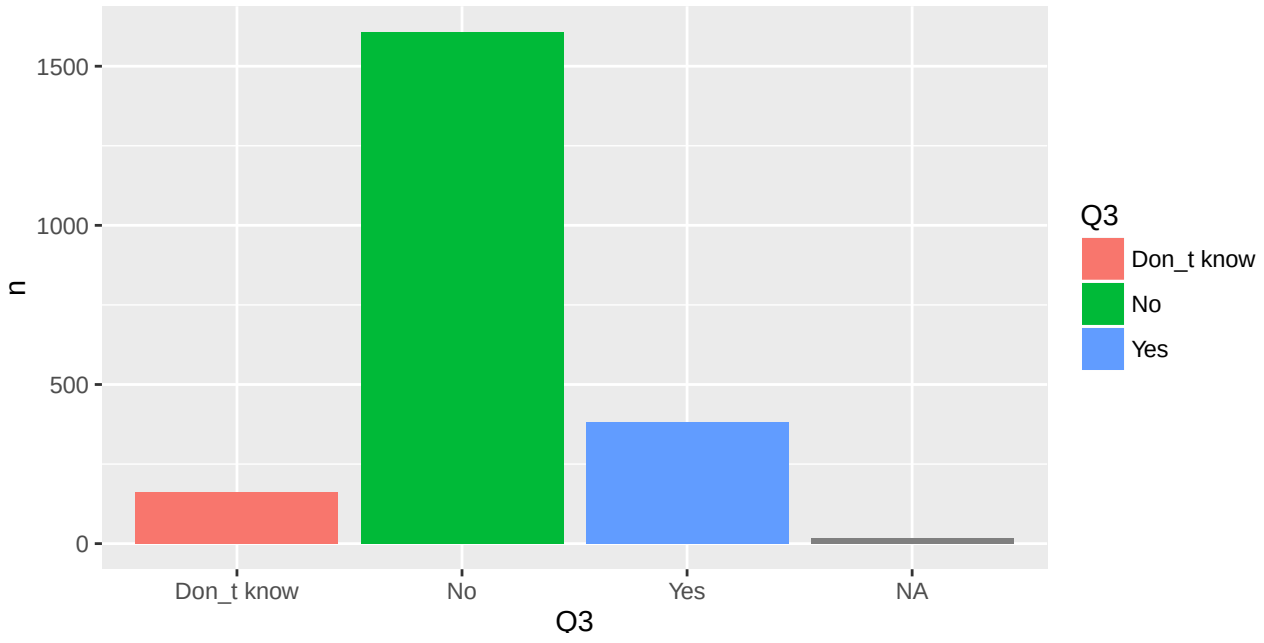


Q3. Has any other person in your household had an illness with influenza like symptoms since August 2015?

```
# all
with(data, table(Q3))
```

```
## Q3
## Don_t know         No        Yes
##        161       1608        383
```

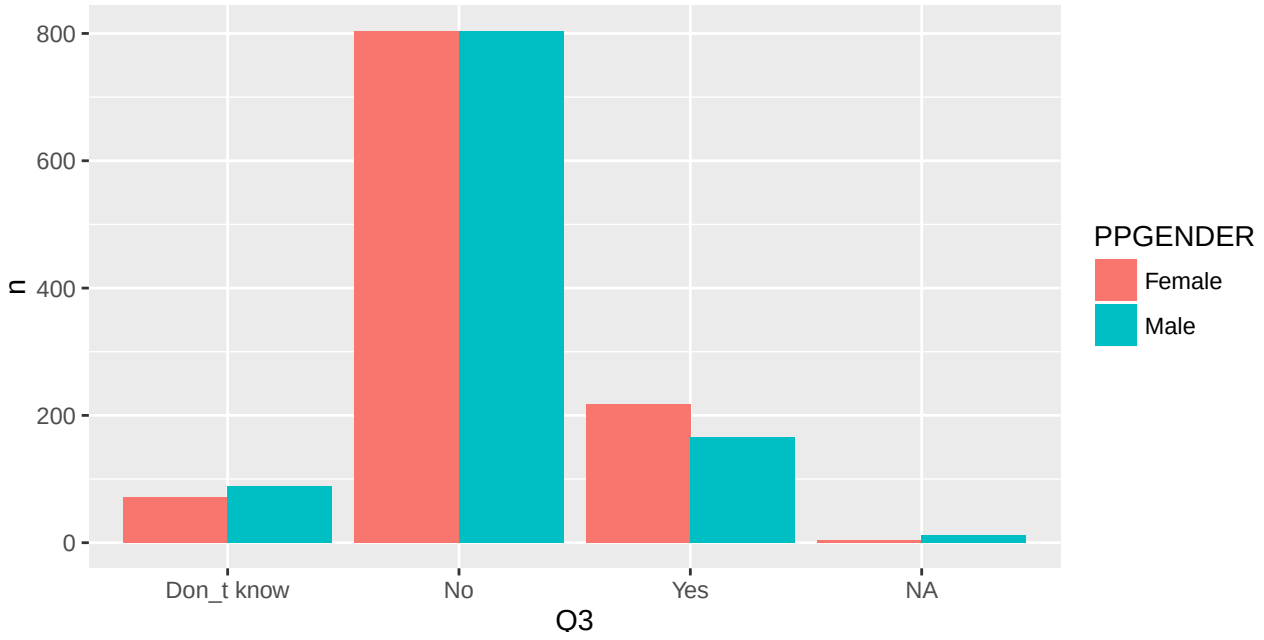
```
q3 <- data %>%
  count(Q3)
ggplot(q3, aes(x = Q3, y = n, fill = Q3)) + geom_bar(stat = 'identity')
```




```
# by gender
with(data, table(Q3, PPGENDER))
```

```
##             PPGENDER
## Q3           Female Male
##   Don_t know     72   89
##   No            804  804
##   Yes           217  166
```

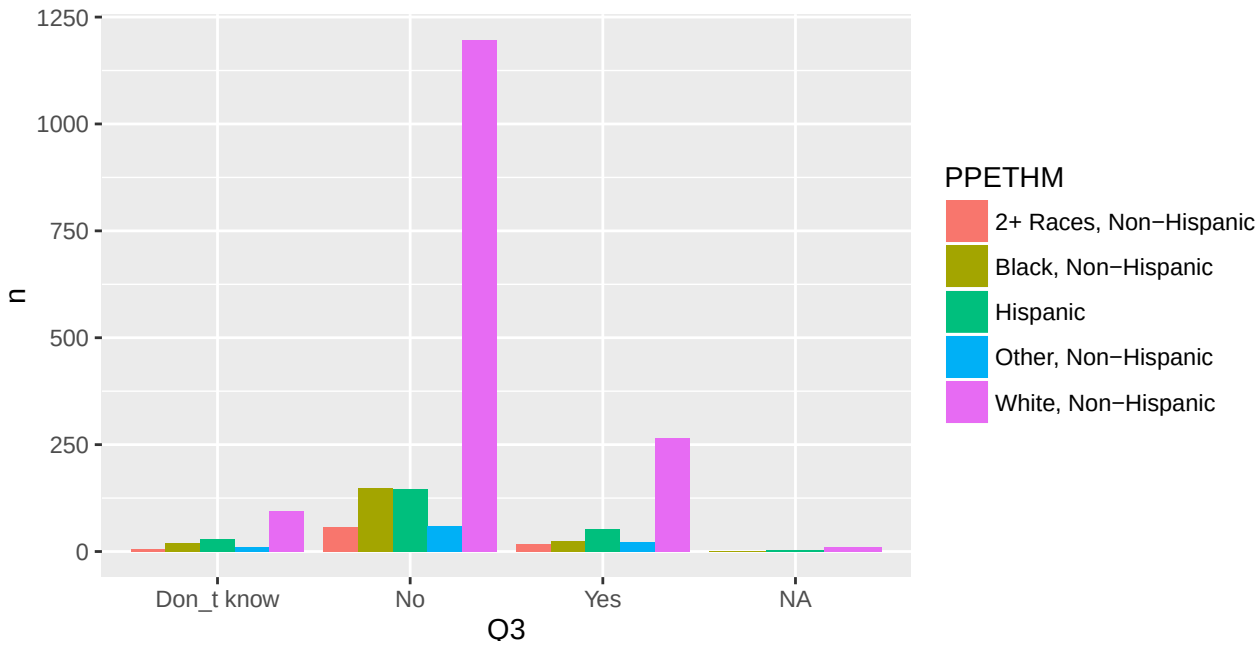
```
q3 <- data %>%
  count(Q3, PPGENDER)
ggplot(q3, aes(x = Q3, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



```r
# by ethnicity
with(data, table(Q3, PPETHM))
```

```
##             PPETHM
## Q3           2+ Races, Non-Hispanic Black, Non-Hispanic Hispanic
##   Don_t know                      6                  19       30
##   No                             57                 149      146
##   Yes                            17                  25       53
##             PPETHM
## Q3           Other, Non-Hispanic White, Non-Hispanic
##   Don_t know                  11                  95
##   No                          59                1197
##   Yes                         23                 265
```

```r
q3 <- data %>%
  count(Q3, PPETHM)
ggplot(q3, aes(x = Q3, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```

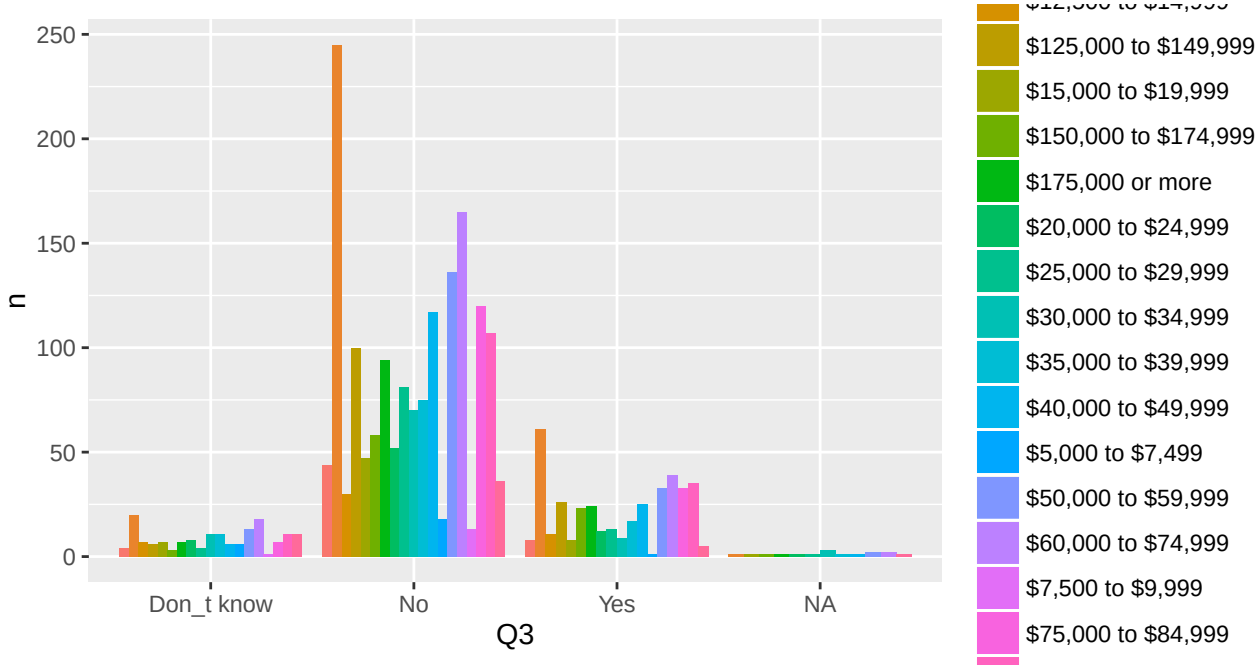



```r
# by income
with(data, table(Q3, PPINCIMP))
```

```
##             PPINCIMP
## Q3           $10,000 to $12,499 $100,000 to $124,999 $12,500 to $14,999
##   Don_t know                  4                  20                  7
##   No                         44                 245                 30
##   Yes                         8                  61                 11
##             PPINCIMP
## Q3           $125,000 to $149,999 $15,000 to $19,999 $150,000 to $174,999
##   Don_t know                    6                  7                   3
##   No                          100                 47                  58
```

```
##   Yes                          26                  8                  23
##             PPINCIMP
## Q3           $175,000 or more $20,000 to $24,999 $25,000 to $29,999
##   Don_t know                7                  8                   4
##   No                       94                 52                  81
##   Yes                      24                 12                  13
##             PPINCIMP
## Q3           $30,000 to $34,999 $35,000 to $39,999 $40,000 to $49,999
##   Don_t know                 11                 11                   6
##   No                         70                 75                 117
##   Yes                         9                 17                  25
##             PPINCIMP
## Q3           $5,000 to $7,499 $50,000 to $59,999 $60,000 to $74,999
##   Don_t know                6                 13                  18
##   No                       18                136                 165
##   Yes                       1                 33                  39
##             PPINCIMP
## Q3           $7,500 to $9,999 $75,000 to $84,999 $85,000 to $99,999
##   Don_t know                1                  7                  11
##   No                       13                120                 107
##   Yes                       0                 33                  35
##             PPINCIMP
## Q3           Less than $5,000
##   Don_t know               11
##   No                       36
##   Yes                       5
```

```
q3 <- data %>%
  count(Q3, PPINCIMP)
ggplot(q3, aes(x = Q3, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



Q4. Does your job require you to have a lot of contact with the public?

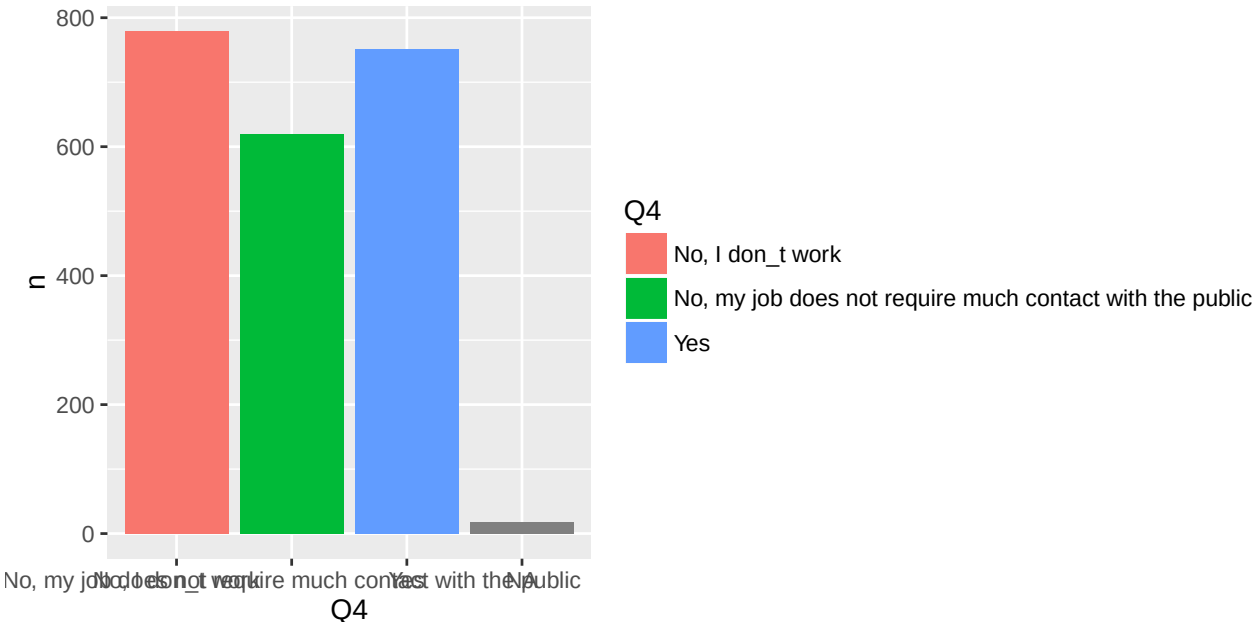
```
# all
with(data, table(Q4))
```

```
## Q4
##                                          No, I don_t work
##                                                       779
## No, my job does not require much contact with the public
##                                                       620
##                                                       Yes
##                                                       751
```

```
(
q4 <- data %>%
  count(Q4)
)
```

```
## Source: local data frame [4 x 2]
##
##                                                       Q4     n
##                                                   <fctr> <int>
## 1                                        No, I don_t work   779
## 2 No, my job does not require much contact with the public   620
## 3                                                    Yes   751
## 4                                                     NA    18
```

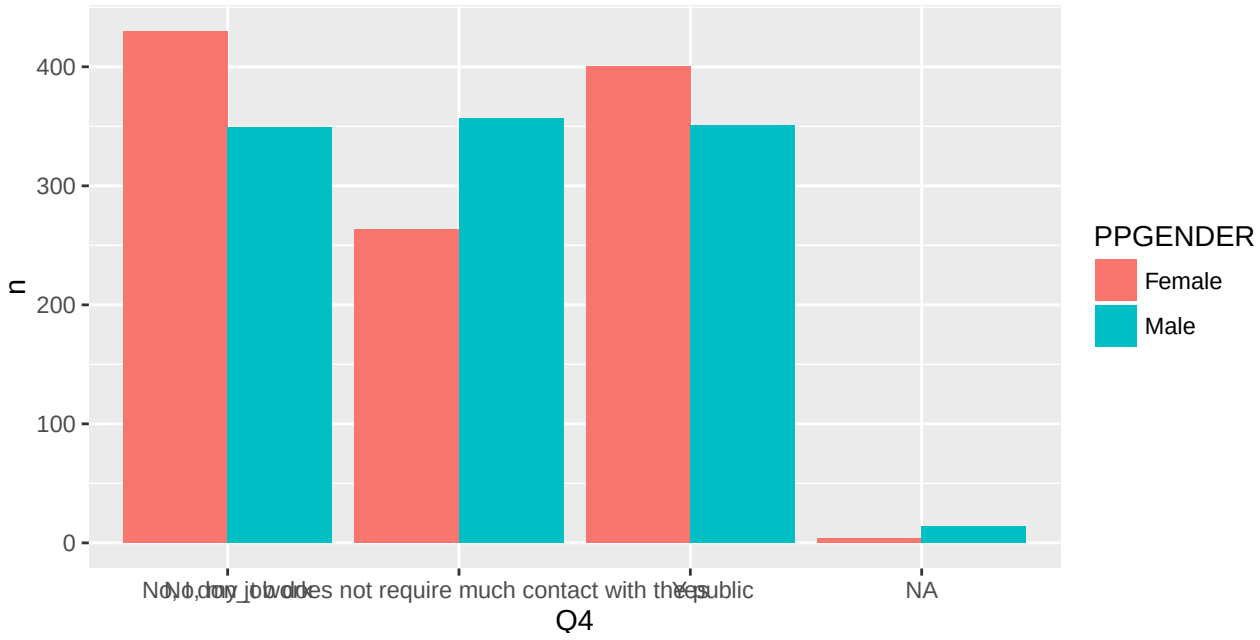
```
ggplot(q4, aes(x = Q4, y = n, fill = Q4)) + geom_bar(stat = 'identity')
```



```r
# by gender
with(data, table(Q4, PPGENDER))
```

```
##                                                                PPGENDER
## Q4                                                              Female Male
##   No, I don_t work                                                 430  349
##   No, my job does not require much contact with the public         263  357
##   Yes                                                              400  351
```

```r
q4 <- data %>%
  count(Q4, PPGENDER)
ggplot(q4, aes(x = Q4, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```

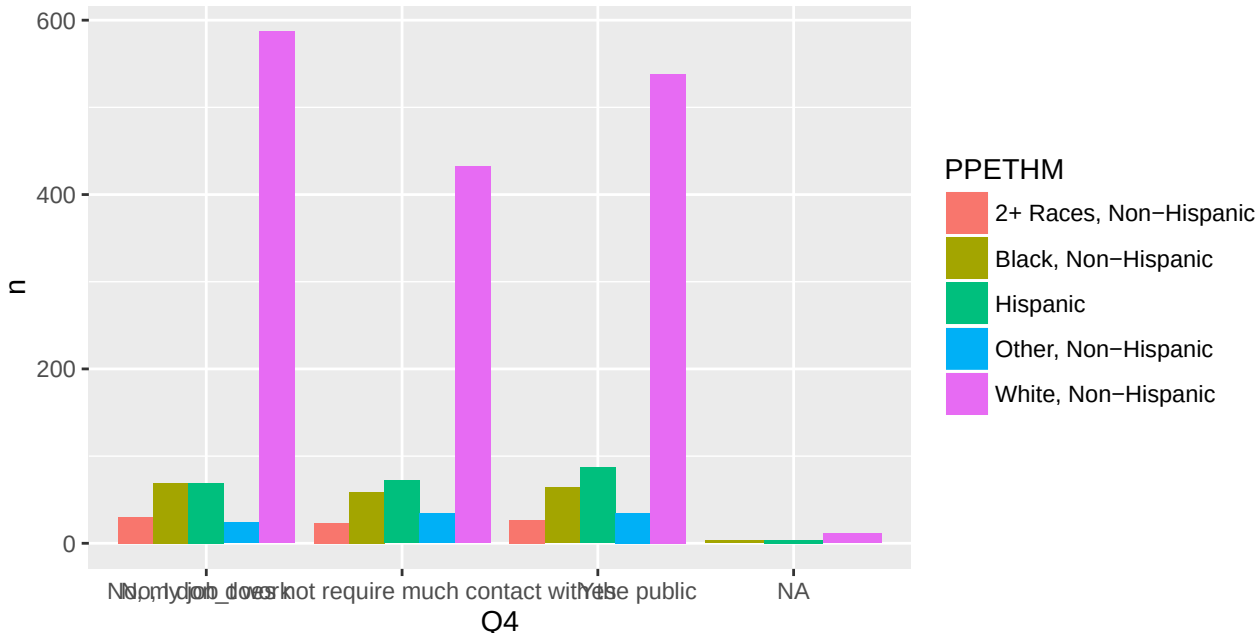



```r
# by ethnicity
with(data, table(Q4, PPETHM))
```

```
##                                                                PPETHM
## Q4                                                              2+ Races, Non-Hispanic
##   No, I don_t work                                                                  30
##   No, my job does not require much contact with the public                          23
##   Yes                                                                               27
##                                                                PPETHM
## Q4                                                              Black, Non-Hispanic
##   No, I don_t work                                                                69
##   No, my job does not require much contact with the public                        59
##   Yes                                                                             64
##                                                                PPETHM
## Q4                                                              Hispanic
##   No, I don_t work                                                    69
##   No, my job does not require much contact with the public            72
```

```
##   Yes                                                                  87
##                                                                  PPETHM
## Q4                                                                Other, Non-Hispanic
##   No, I don_t work                                                                24
##   No, my job does not require much contact with the public                        34
##   Yes                                                                             35
##                                                                  PPETHM
## Q4                                                                White, Non-Hispanic
##   No, I don_t work                                                              587
##   No, my job does not require much contact with the public                      432
##   Yes                                                                           538
```

```
q4 <- data %>%
  count(Q4, PPETHM)
ggplot(q4, aes(x = Q4, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



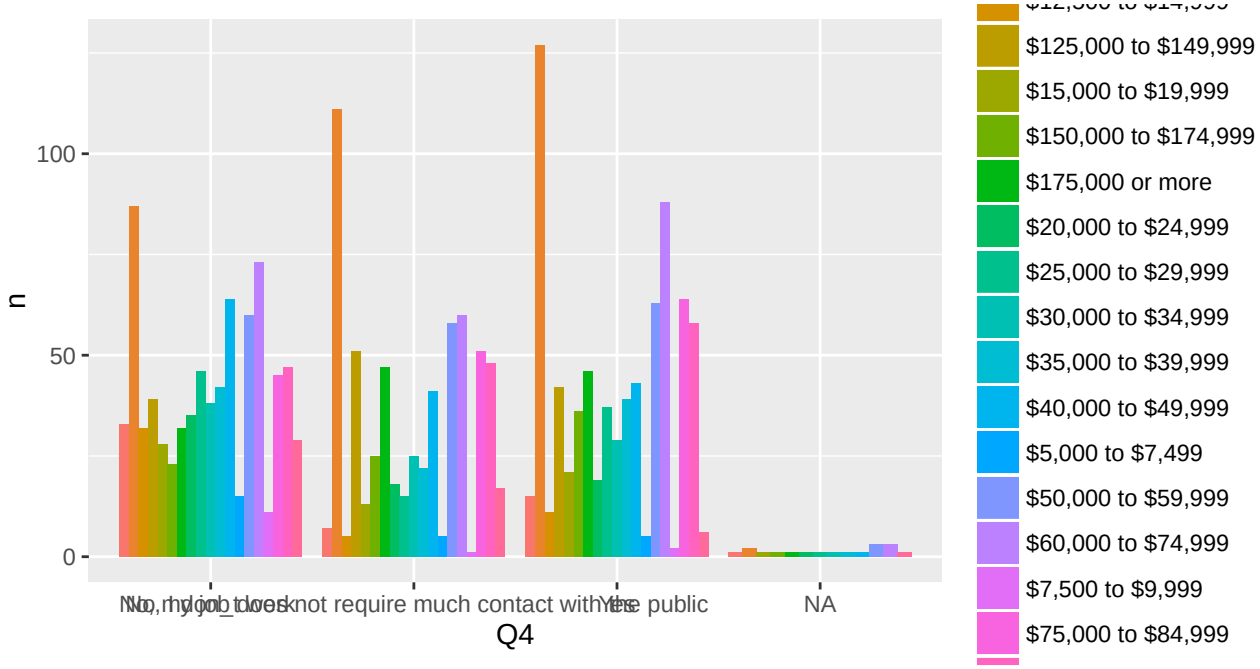

```
# by income
with(data, table(Q4, PPINCIMP))
```

```
##                                                                  PPINCIMP
## Q4                                                                $10,000 to $12,499
##   No, I don_t work                                                              33
##   No, my job does not require much contact with the public                       7
##   Yes                                                                           15
##                                                                  PPINCIMP
## Q4                                                                $100,000 to $124,999
##   No, I don_t work                                                              87
##   No, my job does not require much contact with the public                     111
##   Yes                                                                          127
##                                                                  PPINCIMP
## Q4                                                                $12,500 to $14,999
```

```
##   No, I don_t work                                                              32
##   No, my job does not require much contact with the public                       5
##   Yes                                                                           11
##                                                              PPINCIMP
## Q4                                                            $125,000 to $149,999
##   No, I don_t work                                                              39
##   No, my job does not require much contact with the public                      51
##   Yes                                                                           42
##                                                              PPINCIMP
## Q4                                                            $15,000 to $19,999
##   No, I don_t work                                                              28
##   No, my job does not require much contact with the public                      13
##   Yes                                                                           21
##                                                              PPINCIMP
## Q4                                                            $150,000 to $174,999
##   No, I don_t work                                                              23
##   No, my job does not require much contact with the public                      25
##   Yes                                                                           36
##                                                              PPINCIMP
## Q4                                                            $175,000 or more
##   No, I don_t work                                                            32
##   No, my job does not require much contact with the public                    47
##   Yes                                                                         46
##                                                              PPINCIMP
## Q4                                                            $20,000 to $24,999
##   No, I don_t work                                                              35
##   No, my job does not require much contact with the public                      18
##   Yes                                                                           19
##                                                              PPINCIMP
## Q4                                                            $25,000 to $29,999
##   No, I don_t work                                                              46
##   No, my job does not require much contact with the public                      15
##   Yes                                                                           37
##                                                              PPINCIMP
## Q4                                                            $30,000 to $34,999
##   No, I don_t work                                                              38
##   No, my job does not require much contact with the public                      25
##   Yes                                                                           29
##                                                              PPINCIMP
## Q4                                                            $35,000 to $39,999
##   No, I don_t work                                                              42
##   No, my job does not require much contact with the public                      22
##   Yes                                                                           39
##                                                              PPINCIMP
## Q4                                                            $40,000 to $49,999
##   No, I don_t work                                                              64
##   No, my job does not require much contact with the public                      41
##   Yes                                                                           43
##                                                              PPINCIMP
## Q4                                                            $5,000 to $7,499
##   No, I don_t work                                                            15
##   No, my job does not require much contact with the public                     5
##   Yes                                                                          5
##                                                              PPINCIMP
```

```
## Q4                                                              $50,000 to $59,999
##   No, I don_t work                                                              60
##   No, my job does not require much contact with the public                      58
##   Yes                                                                           63
##                                                                 PPINCIMP
## Q4                                                              $60,000 to $74,999
##   No, I don_t work                                                              73
##   No, my job does not require much contact with the public                      60
##   Yes                                                                           88
##                                                                 PPINCIMP
## Q4                                                              $7,500 to $9,999
##   No, I don_t work                                                             11
##   No, my job does not require much contact with the public                      1
##   Yes                                                                           2
##                                                                 PPINCIMP
## Q4                                                              $75,000 to $84,999
##   No, I don_t work                                                              45
##   No, my job does not require much contact with the public                      51
##   Yes                                                                           64
##                                                                 PPINCIMP
## Q4                                                              $85,000 to $99,999
##   No, I don_t work                                                              47
##   No, my job does not require much contact with the public                      48
##   Yes                                                                           58
##                                                                 PPINCIMP
## Q4                                                              Less than $5,000
##   No, I don_t work                                                             29
##   No, my job does not require much contact with the public                     17
##   Yes                                                                           6
```

```
q4 <- data %>%
  count(Q4, PPINCIMP)
ggplot(q4, aes(x = Q4, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```

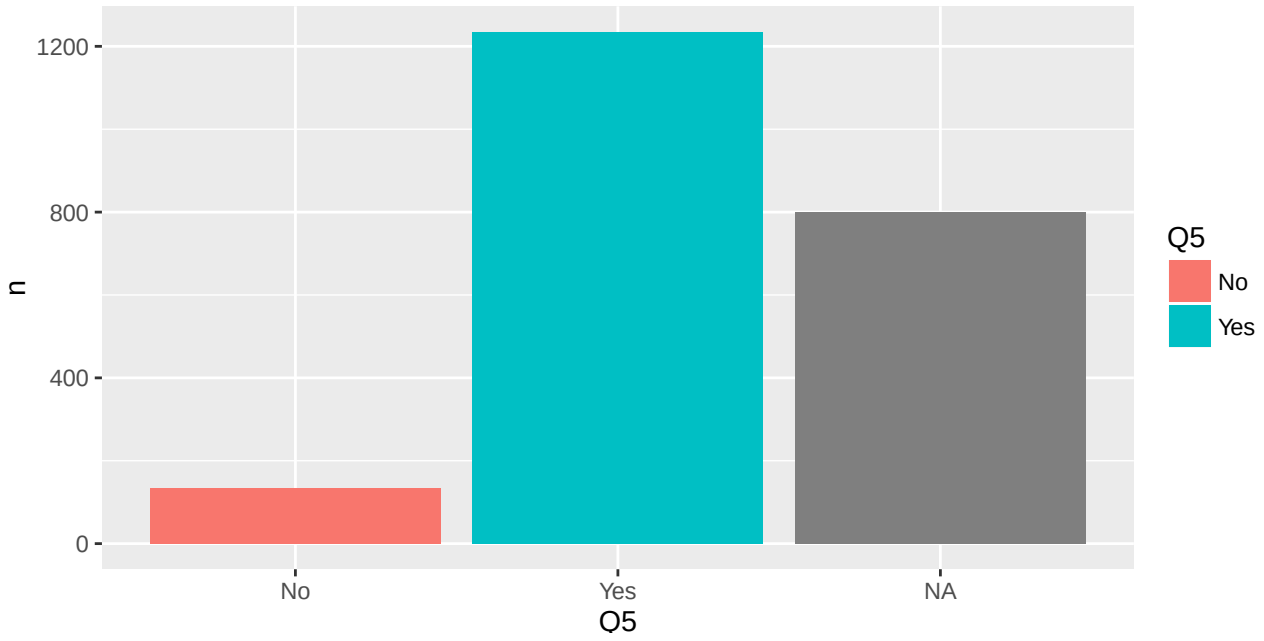


Q5. Do you have a car that you can use to travel to work?

```
# all
with(data, table(Q5))
```

```
## Q5
##   No  Yes
##  133 1235
```

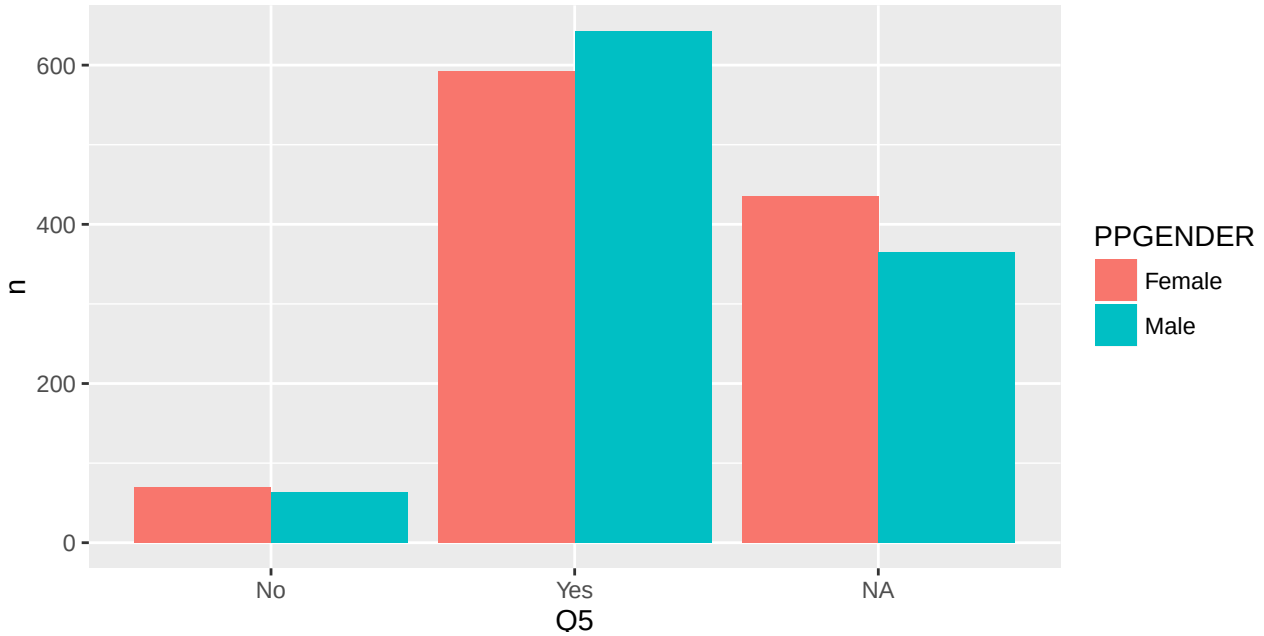
```
q5 <- data %>%
  count(Q5)
ggplot(q5, aes(x = Q5, y = n, fill = Q5)) + geom_bar(stat = 'identity')
```



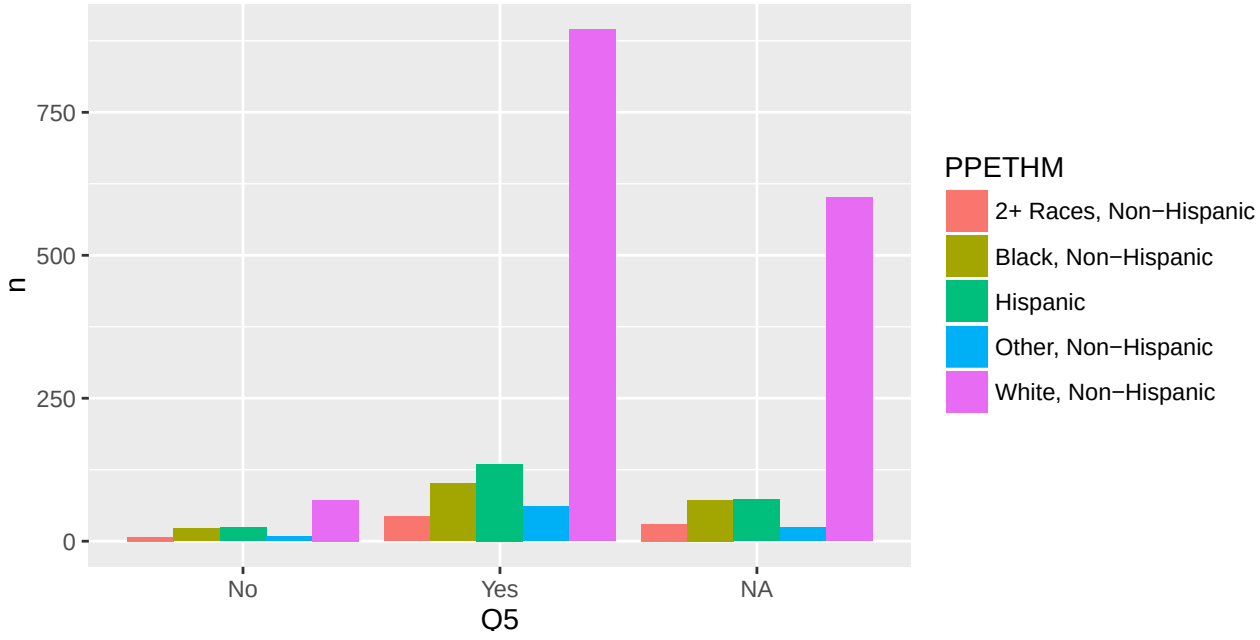

```
# by gender
with(data, table(PPGENDER, Q5))
```

```
##         Q5
## PPGENDER  No Yes
##   Female  70 592
##   Male    63 643
```

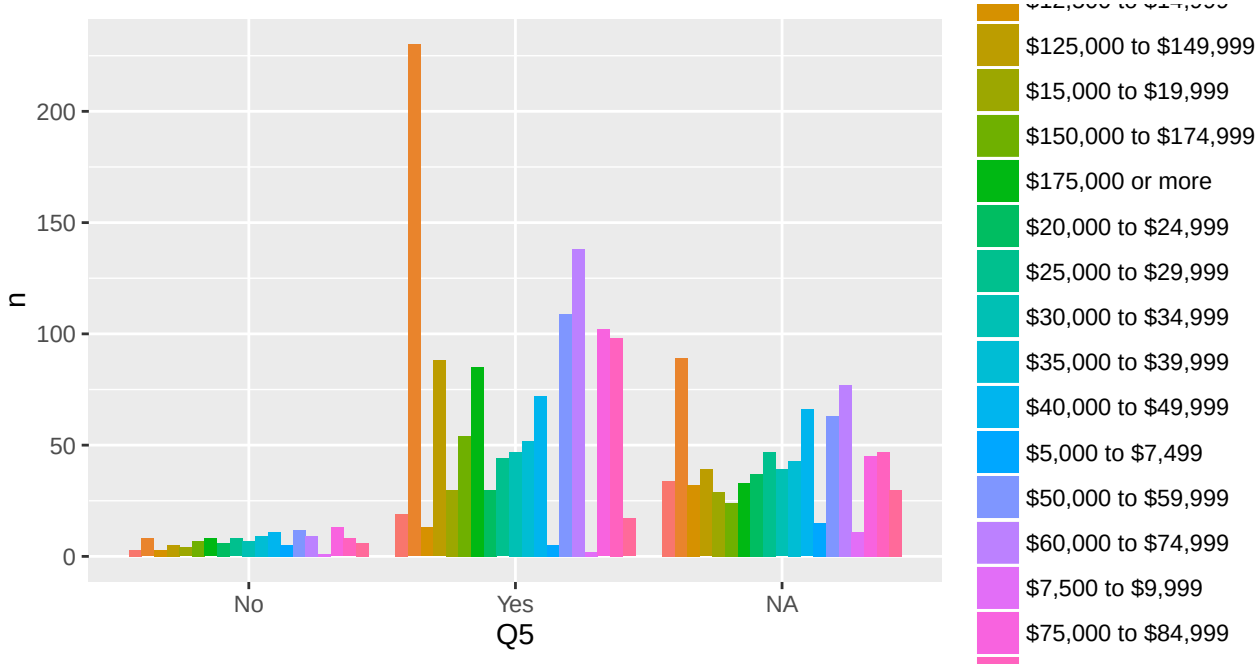
```
q5 <- data %>%
  count(Q5, PPGENDER)
ggplot(q5, aes(x = Q5, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



```r
# by ethnicity
q5 <- data %>%
  count(Q5, PPETHM)
ggplot(q5, aes(x = Q5, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```r
# by income
q5 <- data %>%
  count(Q5, PPINCIMP)
ggplot(q5, aes(x = Q5, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```

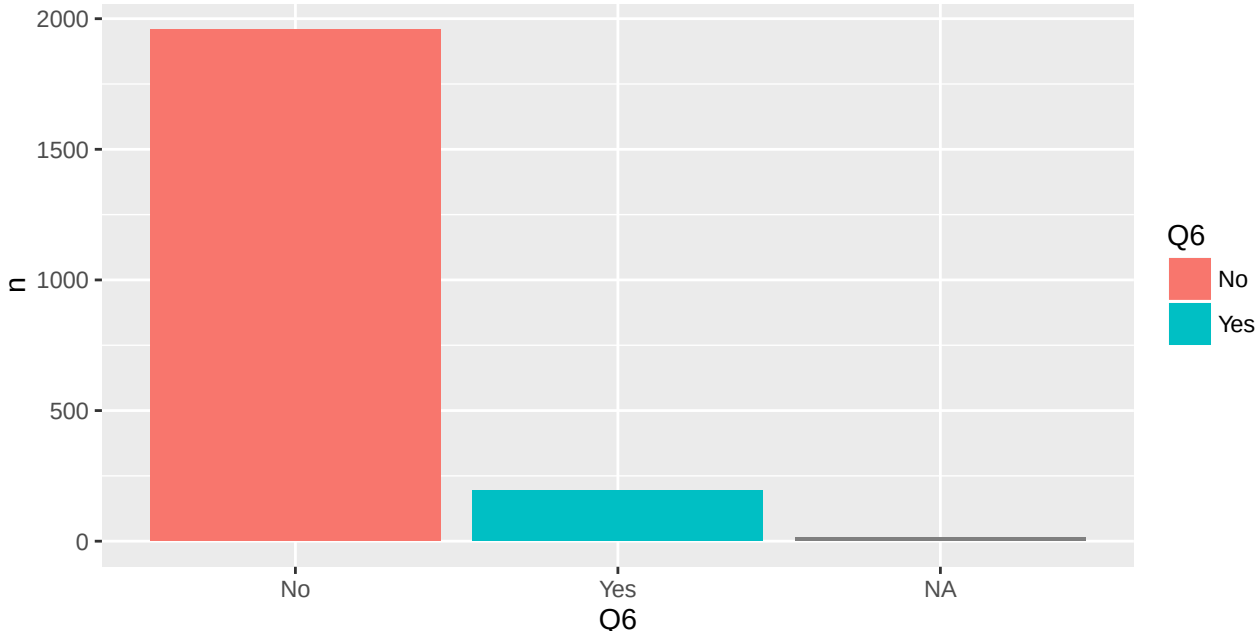


Q6. Do you regularly use public transportation?

```
# all
with(data, table(Q6))
```

```
## Q6
##   No  Yes
## 1959  194
```

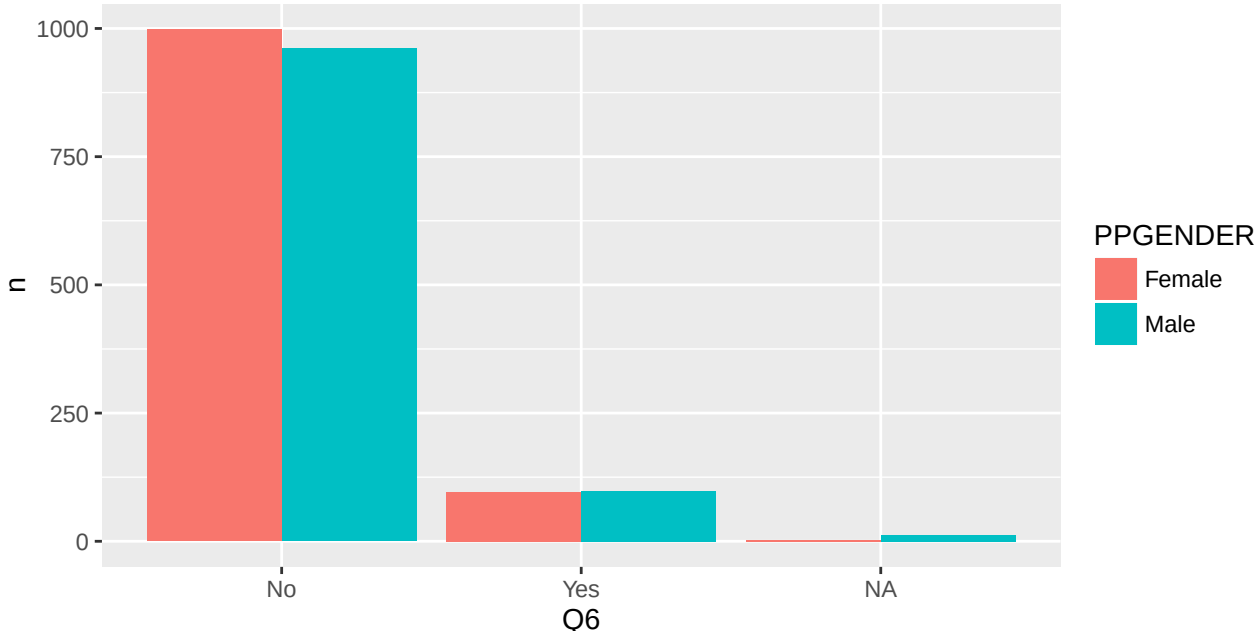
```
q6 <- data %>%
  count(Q6)
ggplot(q6, aes(x = Q6, y = n, fill = Q6)) + geom_bar(stat = 'identity')
```




```
# by gender
# with(data, table(PPGENDER, Q6))
(q6 <- data %>%
  count(Q6, PPGENDER)
)
```

```
## Source: local data frame [6 x 3]
## Groups: Q6 [?]
##
##       Q6 PPGENDER     n
##   (fctr)   (fctr) (int)
## 1     No   Female   998
## 2     No     Male   961
## 3    Yes   Female    96
## 4    Yes     Male    98
## 5     NA   Female     3
## 6     NA     Male    12
```

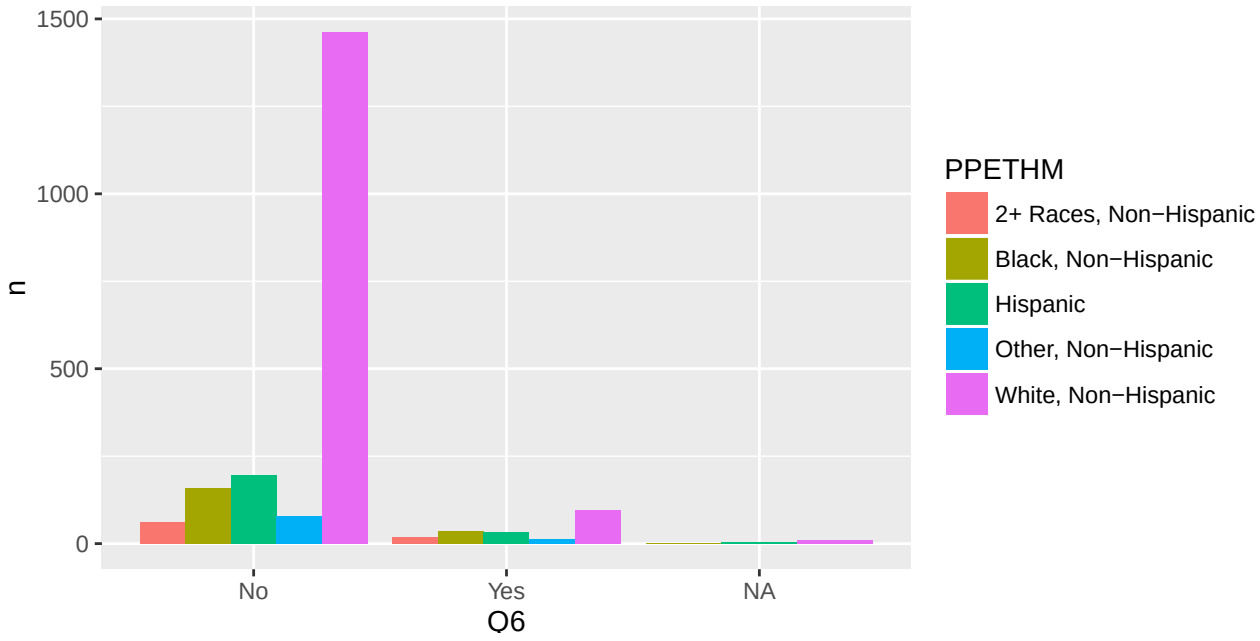
```
ggplot(q6, aes(x = Q6, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# by ethnicity
(q6 <- data %>%
  count(Q6, PPETHM)
)
```

```
## Source: local data frame [13 x 3]
## Groups: Q6 [?]
##
##        Q6                PPETHM     n
##    (fctr)                (fctr) (int)
## 1      No 2+ Races, Non-Hispanic    62
## 2      No    Black, Non-Hispanic   158
## 3      No              Hispanic   196
## 4      No    Other, Non-Hispanic    80
## 5      No    White, Non-Hispanic  1463
## 6     Yes 2+ Races, Non-Hispanic    18
## 7     Yes    Black, Non-Hispanic    36
## 8     Yes              Hispanic    32
## 9     Yes    Other, Non-Hispanic    13
## 10    Yes    White, Non-Hispanic    95
## 11     NA    Black, Non-Hispanic     1
## 12     NA              Hispanic     4
## 13     NA    White, Non-Hispanic    10
```

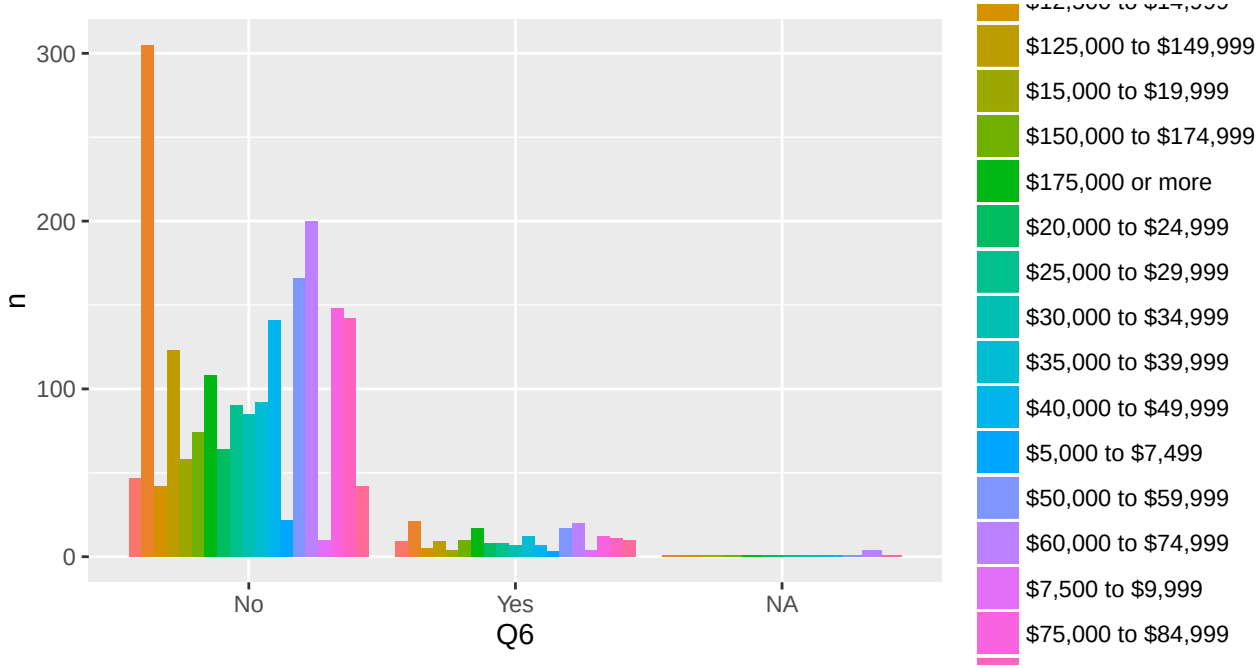
```
ggplot(q6, aes(x = Q6, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge())
```




```
# by income
(q6 <- data %>%
  count(Q6, PPINCIMP)
)
```

```
## Source: local data frame [50 x 3]
## Groups: Q6 [?]
##
##        Q6             PPINCIMP     n
##    (fctr)               (fctr) (int)
## 1      No   $10,000 to $12,499    47
## 2      No $100,000 to $124,999   305
## 3      No   $12,500 to $14,999    42
## 4      No $125,000 to $149,999   123
## 5      No   $15,000 to $19,999    58
## 6      No $150,000 to $174,999    74
## 7      No     $175,000 or more   108
## 8      No   $20,000 to $24,999    64
## 9      No   $25,000 to $29,999    90
## 10     No   $30,000 to $34,999    85
## ..    ...                  ...   ...
```

```
ggplot(q6, aes(x = Q6, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge())
```



Q7. What types of public transportation do you regularly use?

```
Q7 <- data2 %>%
  select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q7_1_Bus:Q7_otherText) %>%
  gather("q", "r", Q7_1_Bus:Q7_7_Other)

# Q7
with(Q7, table(q, r))
```

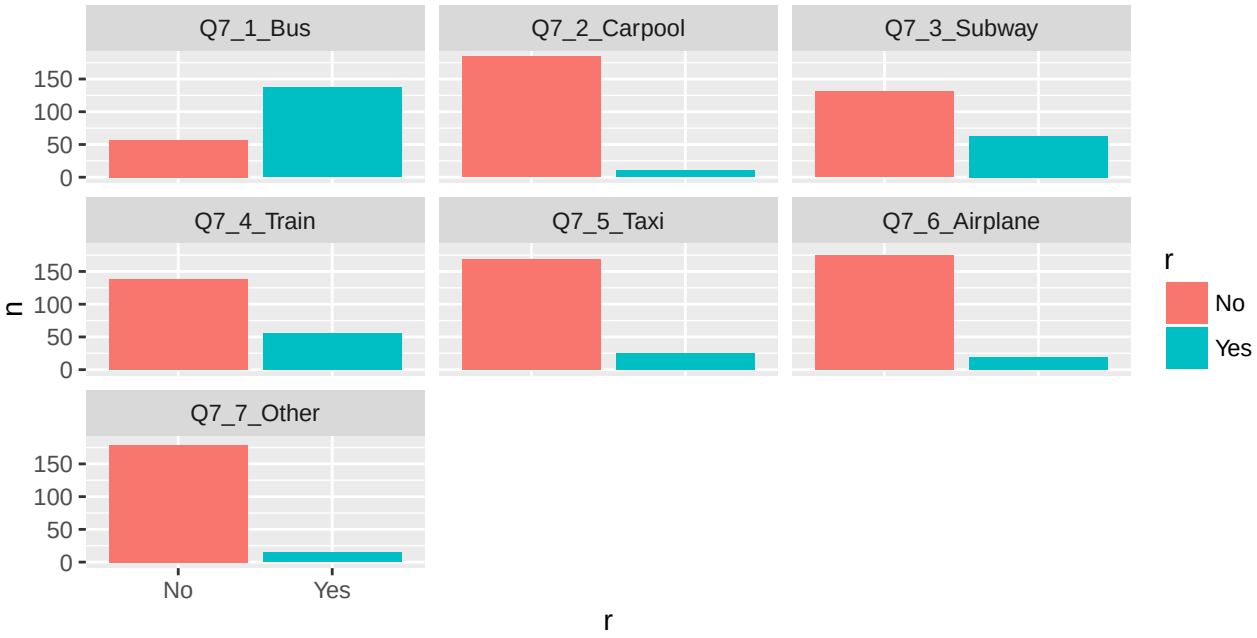
```
##                r
## q                No Yes
##   Q7_1_Bus       57 137
##   Q7_2_Carpool  184  10
##   Q7_3_Subway   131  63
##   Q7_4_Train    139  55
##   Q7_5_Taxi     169  25
##   Q7_6_Airplane 175  19
##   Q7_7_Other    179  15
```

```
(
q7 <- Q7 %>%
  count(q, r)
)
```

```
## Source: local data frame [21 x 3]
## Groups: q [?]
##
##               q     r     n
##           (chr) (chr) (int)
```

```
## 1       Q7_1_Bus    No    57
## 2       Q7_1_Bus   Yes   137
## 3       Q7_1_Bus    NA  1974
## 4  Q7_2_Carpool    No   184
## 5  Q7_2_Carpool   Yes    10
## 6  Q7_2_Carpool    NA  1974
## 7   Q7_3_Subway    No   131
## 8   Q7_3_Subway   Yes    63
## 9   Q7_3_Subway    NA  1974
## 10   Q7_4_Train    No   139
## ..           ...   ...   ...
```

```
ggplot(q7[!is.na(q7$r), ], aes(x = r, y = n, fill = r)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```

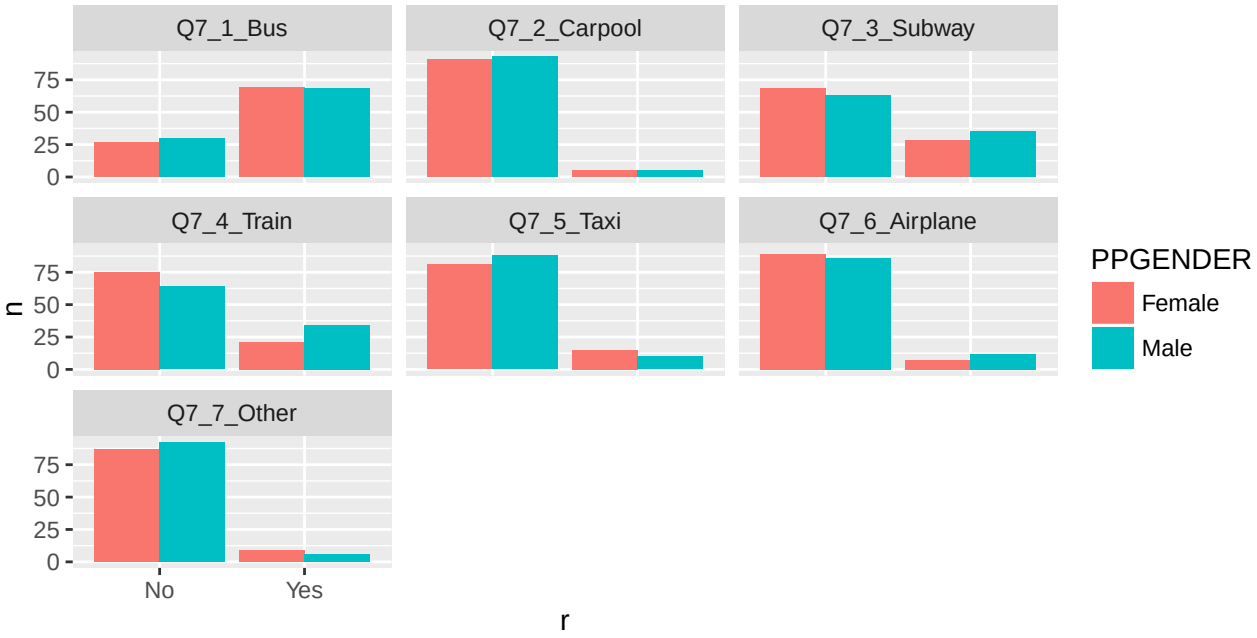



```
# by gender
# with(Q7, table(PPGENDER, r, q))
(q7 <- Q7 %>%
  group_by(PPGENDER, q, r) %>%
  count(PPGENDER, q, r)
)
```

```
## Source: local data frame [42 x 4]
## Groups: PPGENDER, q [?]
##
##    PPGENDER            q     r     n
##      (fctr)        (chr) (chr) (int)
## 1    Female     Q7_1_Bus    No    27
## 2    Female     Q7_1_Bus   Yes    69
## 3    Female     Q7_1_Bus    NA  1001
## 4    Female Q7_2_Carpool    No    91
## 5    Female Q7_2_Carpool   Yes     5
```

```
## 6    Female Q7_2_Carpool    NA  1001
## 7    Female  Q7_3_Subway    No    68
## 8    Female  Q7_3_Subway   Yes    28
## 9    Female  Q7_3_Subway    NA  1001
## 10   Female   Q7_4_Train    No    75
## ..      ...          ...   ...   ...
```

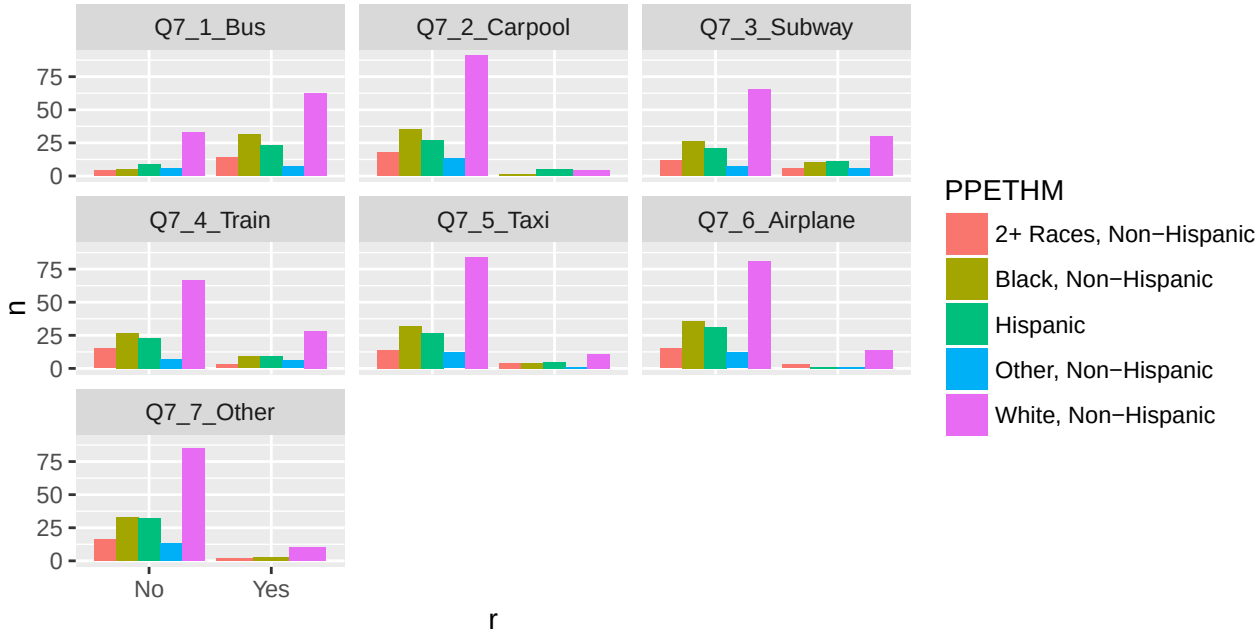
```
ggplot(q7[!is.na(q7$r), ], aes(x = r, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```




```
# by ethnicity
# with(Q7, table(PPETHM, r, q))
(q7 <- Q7 %>%
  group_by(PPETHM, q, r) %>%
  count(PPETHM, q, r)
)
```

```
## Source: local data frame [100 x 4]
## Groups: PPETHM, q [?]
##
##                    PPETHM           q     r     n
##                    (fctr)       (chr) (chr) (int)
## 1  2+ Races, Non-Hispanic     Q7_1_Bus    No     4
## 2  2+ Races, Non-Hispanic     Q7_1_Bus   Yes    14
## 3  2+ Races, Non-Hispanic     Q7_1_Bus    NA    62
## 4  2+ Races, Non-Hispanic Q7_2_Carpool    No    18
## 5  2+ Races, Non-Hispanic Q7_2_Carpool    NA    62
## 6  2+ Races, Non-Hispanic  Q7_3_Subway    No    12
## 7  2+ Races, Non-Hispanic  Q7_3_Subway   Yes     6
## 8  2+ Races, Non-Hispanic  Q7_3_Subway    NA    62
## 9  2+ Races, Non-Hispanic   Q7_4_Train    No    15
## 10 2+ Races, Non-Hispanic   Q7_4_Train   Yes     3
## ..                    ...          ...   ...   ...
```

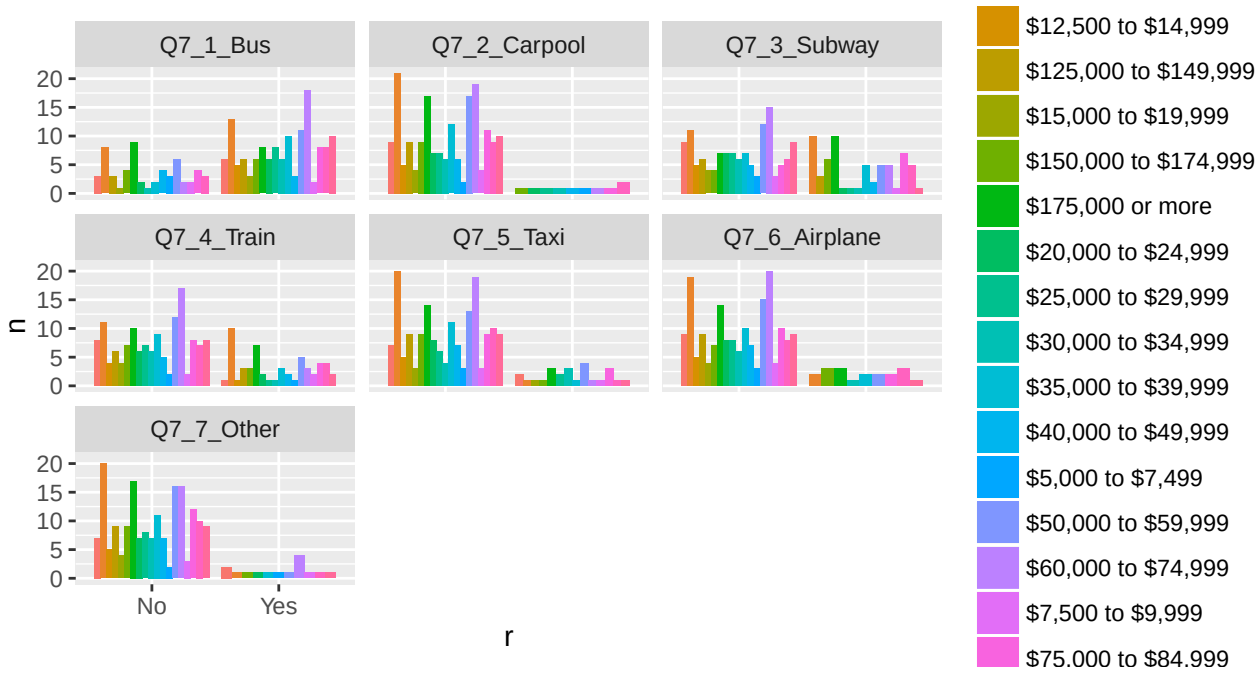
```
ggplot(q7[!is.na(q7$r), ], aes(x = r, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```




```
# by income
# with(Q7, table(q, r, PPINCIMP))
(q7 <- Q7 %>%
  group_by(PPINCIMP, q, r) %>%
  count(PPINCIMP, q, r)
)
```

```
## Source: local data frame [357 x 4]
## Groups: PPINCIMP, q [?]
##
##              PPINCIMP            q     r     n
##                 (fctr)        (chr) (chr) (int)
## 1  $10,000 to $12,499     Q7_1_Bus    No     3
## 2  $10,000 to $12,499     Q7_1_Bus   Yes     6
## 3  $10,000 to $12,499     Q7_1_Bus    NA    47
## 4  $10,000 to $12,499 Q7_2_Carpool    No     9
## 5  $10,000 to $12,499 Q7_2_Carpool    NA    47
## 6  $10,000 to $12,499  Q7_3_Subway    No     9
## 7  $10,000 to $12,499  Q7_3_Subway    NA    47
## 8  $10,000 to $12,499   Q7_4_Train    No     8
## 9  $10,000 to $12,499   Q7_4_Train   Yes     1
## 10 $10,000 to $12,499   Q7_4_Train    NA    47
## ..                ...          ...   ...   ...
```

```
ggplot(q7[!is.na(q7$r), ], aes(x = r, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```



Q8. For what types of activities do you regularly use public transportation?

```
Q8 <- data2 %>%
  select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q8_1_Work:Q8_otherText) %>%
  gather("q", "r", Q8_1_Work:Q8_6_Other)
```

Q9. Do other members of your household regularly use public transportation?

```
with(data, table(Q9))
```

```
## Q9
## Don_t know         No        Yes
##         32       1935        183
```

Q10. What types of public transportation do other members of your household regularly use?

Q10 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q11. How do you rate your risk of getting influenza if you visited each of the following locations?

```
Q11 <- data2 %>%
  select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q11_1_Work:Q11_OtherText_Codes) %>%
  gather("q", "r", Q11_1_Work:Q11_11_Other)


# all
with(Q11, table(q, r))
```

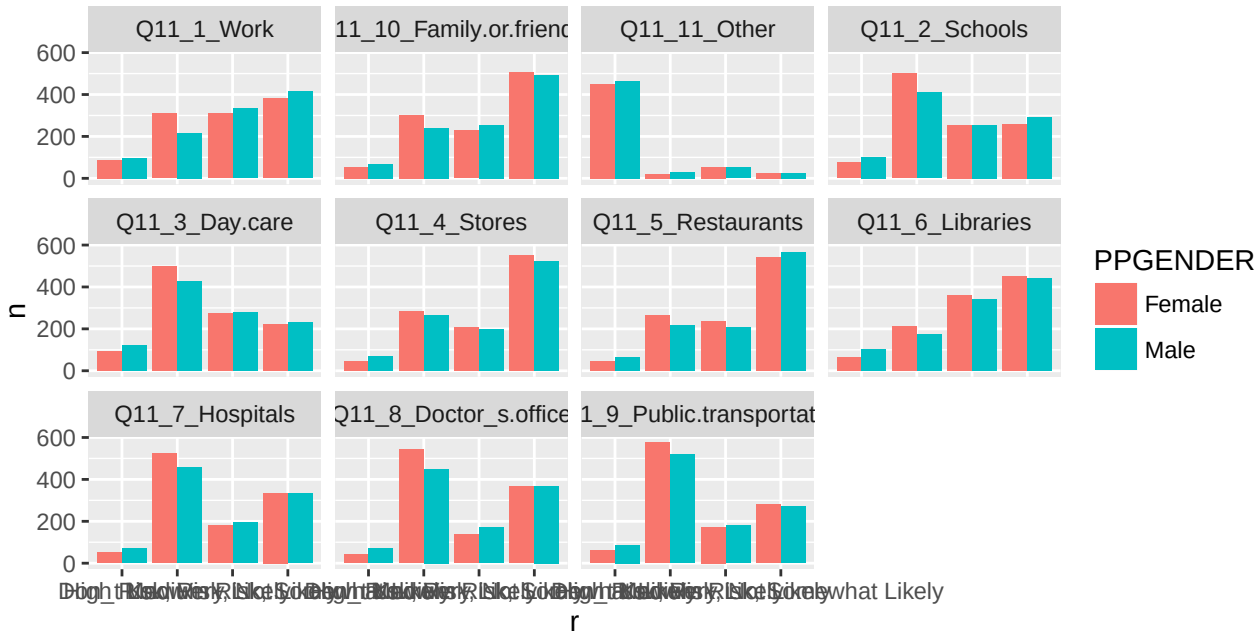
```
##                               r
## q                              Don_t Know High Risk, Very Likely
##   Q11_1_Work                          185                    524
##   Q11_10_Family.or.friends            121                    541
##   Q11_11_Other                        915                     51
##   Q11_2_Schools                       178                    909
##   Q11_3_Day.care                      214                    924
##   Q11_4_Stores                        115                    551
##   Q11_5_Restaurants                   111                    483
##   Q11_6_Libraries                     169                    386
##   Q11_7_Hospitals                     123                    982
##   Q11_8_Doctor_s.office               110                    994
##   Q11_9_Public.transportation         147                   1093
##                               r
## q                              Low Risk, Not Likely
##   Q11_1_Work                                    643
##   Q11_10_Family.or.friends                      485
##   Q11_11_Other                                  104
##   Q11_2_Schools                                 508
##   Q11_3_Day.care                                554
##   Q11_4_Stores                                  405
##   Q11_5_Restaurants                             442
##   Q11_6_Libraries                               700
##   Q11_7_Hospitals                               374
##   Q11_8_Doctor_s.office                         308
##   Q11_9_Public.transportation                   353
##                               r
## q                              Medium Risk, Somewhat Likely
##   Q11_1_Work                                          795
##   Q11_10_Family.or.friends                           1000
##   Q11_11_Other                                         54
##   Q11_2_Schools                                       551
##   Q11_3_Day.care                                      454
##   Q11_4_Stores                                       1076
##   Q11_5_Restaurants                                  1111
##   Q11_6_Libraries                                     890
##   Q11_7_Hospitals                                     669
##   Q11_8_Doctor_s.office                               733
##   Q11_9_Public.transportation                         551
```

```
q11 <- Q11 %>%
  group_by(q, r) %>%
  count(q, r)
ggplot(q11[!is.na(q11$r), ], aes(x = r, y = n, fill = r)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```


```
# by gender
# with(Q7, table(PPGENDER, r, q))
(q11 <- Q11 %>%
  group_by(PPGENDER, q, r) %>%
  count(PPGENDER, q, r)
)
```

```
## Source: local data frame [110 x 4]
## Groups: PPGENDER, q [?]
##
##    PPGENDER                         q                            r     n
##      (fctr)                     (chr)                        (chr) (int)
## 1    Female                Q11_1_Work                   Don_t Know    89
## 2    Female                Q11_1_Work       High Risk, Very Likely   309
## 3    Female                Q11_1_Work         Low Risk, Not Likely   310
## 4    Female                Q11_1_Work Medium Risk, Somewhat Likely   381
## 5    Female                Q11_1_Work                           NA     8
## 6    Female Q11_10_Family.or.friends                   Don_t Know    53
## 7    Female Q11_10_Family.or.friends       High Risk, Very Likely   302
## 8    Female Q11_10_Family.or.friends         Low Risk, Not Likely   229
## 9    Female Q11_10_Family.or.friends Medium Risk, Somewhat Likely   506
## 10   Female Q11_10_Family.or.friends                           NA     7
## ..      ...                      ...                          ...   ...
```

```
ggplot(q11[!is.na(q11$r), ], aes(x = r, y = n, fill = PPGENDER)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```




```
# by ethnicity
# with(Q7, table(PPETHM, r, q))
(q11 <- Q11 %>%
  group_by(PPETHM, q, r) %>%
  count(PPETHM, q, r)
)
```

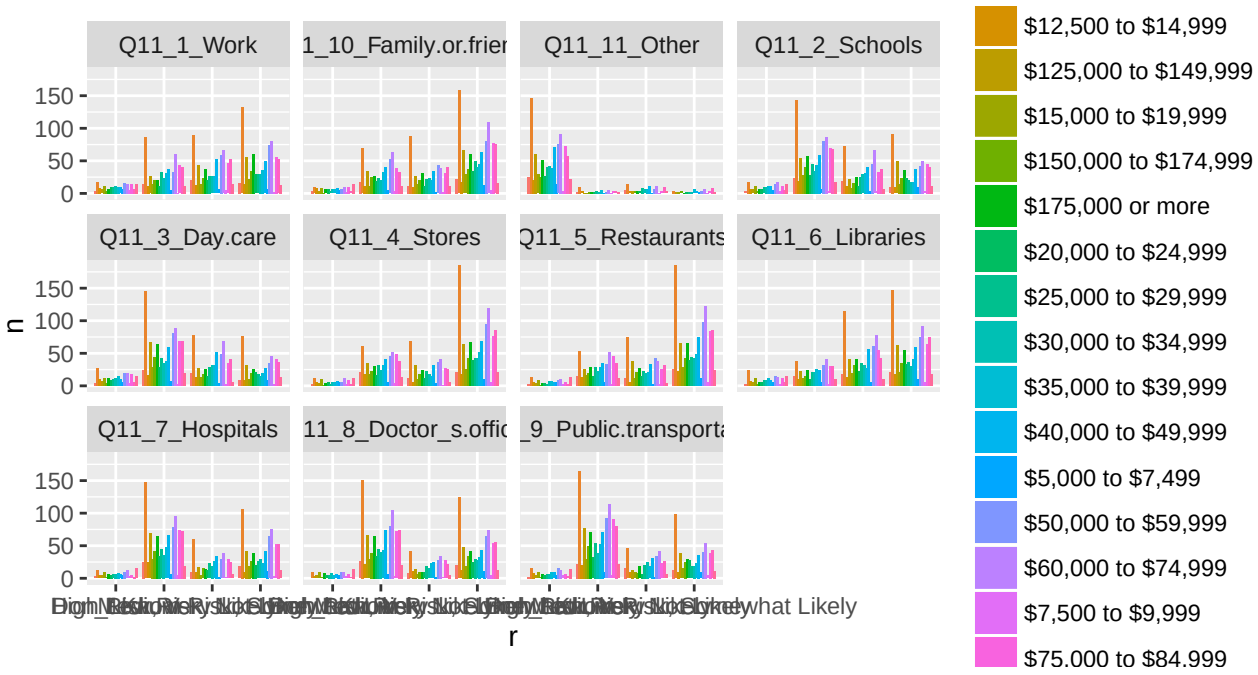
```
## Source: local data frame [275 x 4]
## Groups: PPETHM, q [?]
##
##                    PPETHM                      q
##                    (fctr)                  (chr)
## 1  2+ Races, Non-Hispanic          Q11_1_Work
## 2  2+ Races, Non-Hispanic          Q11_1_Work
## 3  2+ Races, Non-Hispanic          Q11_1_Work
## 4  2+ Races, Non-Hispanic          Q11_1_Work
## 5  2+ Races, Non-Hispanic          Q11_1_Work
## 6  2+ Races, Non-Hispanic Q11_10_Family.or.friends
## 7  2+ Races, Non-Hispanic Q11_10_Family.or.friends
## 8  2+ Races, Non-Hispanic Q11_10_Family.or.friends
## 9  2+ Races, Non-Hispanic Q11_10_Family.or.friends
## 10 2+ Races, Non-Hispanic Q11_10_Family.or.friends
## ..                   ...                     ...
## Variables not shown: r (chr), n (int)
```

```
ggplot(q11[!is.na(q11$r), ], aes(x = r, y = n, fill = PPETHM)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```


```
# by income
# with(Q7, table(q, r, PPINCIMP))
(q11 <- Q11 %>%
  group_by(PPINCIMP, q, r) %>%
  count(PPINCIMP, q, r)
)
```

```
## Source: local data frame [985 x 4]
## Groups: PPINCIMP, q [?]
##
##              PPINCIMP                       q
##                 (fctr)                   (chr)
## 1  $10,000 to $12,499               Q11_1_Work
## 2  $10,000 to $12,499               Q11_1_Work
## 3  $10,000 to $12,499               Q11_1_Work
## 4  $10,000 to $12,499               Q11_1_Work
## 5  $10,000 to $12,499               Q11_1_Work
## 6  $10,000 to $12,499 Q11_10_Family.or.friends
## 7  $10,000 to $12,499 Q11_10_Family.or.friends
## 8  $10,000 to $12,499 Q11_10_Family.or.friends
## 9  $10,000 to $12,499 Q11_10_Family.or.friends
## 10 $10,000 to $12,499 Q11_10_Family.or.friends
## ..                ...                     ...
## Variables not shown: r (chr), n (int)
```

```
ggplot(q11[!is.na(q11$r), ], aes(x = r, y = n, fill = PPINCIMP)) +
  geom_bar(stat = 'identity', position = position_dodge()) + facet_wrap(~q)
```



Q12. Which of the following actions do you take to avoid getting sick?

```
Q12 <- data2 %>%
  select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, 75:91) %>%
  gather("q", "r", Q12_1_Avoid.touching.my.eyes:Q12_15_Other)
```

Q13. Do you get the flu vaccine?

```
with(data, table(Q13))
```

```
## Q13
##       No, never Yes, every year Yes, some years
##             819             908             423
```

Q14. How much do you pay to get an influenza vaccine?

```
with(data, table(Q14))
```

```
## Q14
##            $0    $30 to $60    Don_t know Less than $30 More than $60
##           970            54            80           222             4
```

```
# by gender
with(data, by(Q14, PPGENDER, summary))
```

```
## PPGENDER: Female
##            $0    $30 to $60    Don_t know Less than $30 More than $60
##           514            28            41           101             2
```

```
##          NA's
##           411
## ------------------------------------------------------
## PPGENDER: Male
##            $0    $30 to $60    Don_t know Less than $30 More than $60
##           456            26            39           121             2
##          NA's
##           427
```

Q15. Are you more likely to get a vaccine if others around you get a vaccine?

```
with(data, table(Q15))
```

```
## Q15
##  No, less likely    No, no effect Yes, more likely
##               70              878              381
```

Q16. Are you more likely to get a vaccine if others around you do not get a vaccine?

```
with(data, table(Q16))
```

```
## Q16
##  No, less likely    No, no effect Yes, more likely
##              101              904              313
```

Q17. Do you get a vaccine to protect yourself, protect others, or protect yourself and others?

```
with(data, table(Q17))
```

```
## Q17
##            Protect myself Protect myself and others
##                       381                       921
##            Protect others
##                        22
```

Q18. What are the reasons you would not get an influenza vaccine?

Q18 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q19. Do you have health insurance?

```
with(data, table(Q19))
```

```
## Q19
##   No  Yes
##  154 1994
```

Q20. How effective do you think the influenza vaccine is in protecting people from becoming sick with influenza?

```
with(data, table(Q20))
```

```
## Q20
##                      Don_t know It varies from season to season
##                             228                             433
##                   Not effective             Somewhat effective
##                             144                             961
##                  Very effective
##                             383
```

Q21. Are influenza vaccines covered by your health insurance?

```
with(data, table(Q21))
```

```
## Q21
##                              Don_t know
##                                     500
##                                      No
##                                      55
## Yes, but only part of the cost is paid
##                                     153
##             Yes, the full cost is paid
##                                    1282
```

Q22. Do you do any of the following when you have influenza symptoms?

Q22 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q23. Which of the following actions do you take when you have influenza symptoms to avoid someone else from getting sick?

Q23 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q24. What sources of information do you recall hearing or seeing about influenza outbreaks?

Q24 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q25. If you received information from the news, internet or other public media that there was an influenza outbreak in your community would you do any of the following?

Q25 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q26. Does your household have children?

```
with(data, table(Q26))
```

```
## Q26
##   No  Yes
## 1570  576
```

Q27. What actions do you take when a child in your household has influenza symptoms?

Q27 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q28. Are you a single parent?

```
with(data, table(Q28))
```

```
## Q28
##  No Yes
## 490  86
```

Q29. How do you care for a sick child?

Q29 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q30. How do you care for a sick child?

Q30 <- data2 %>% select(PPGENDER, PPAGE, PPEDUC, PPETHM, PPINCIMP, PPWORK, Q:Q) %>% gather("q", "r", Q:Q)

Q31. How many hours of screen time (time spent watching television, a computer, smartphone, iPad, etc.) do you spend each day on average when you are not sick? Enter 0 if none

```
with(data, summary(Q31))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   2.000   4.000   4.868   6.000  24.000      52
```

```
# by gender
with(data, by(Q31, PPGENDER, summary))
```

```
## PPGENDER: Female
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   2.000   4.000   4.838   6.000  21.000      21
## ------------------------------------------------------------
## PPGENDER: Male
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   2.000   4.000   4.898   6.000  24.000      31
```

Q32. How many hours of screen time do you spend each day on average when you are sick? Enter 0 if none

```
with(data, summary(Q32))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   4.000   4.267   6.000  24.000      61
```

```
# by gender
with(data, by(Q33, PPGENDER, summary))
```

```
## PPGENDER: Female
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   2.000   2.567   3.000   9.000       8
## ------------------------------------------------------------
## PPGENDER: Male
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   2.000   2.594   3.000  14.000      20
```

Q33. How many people, including yourself, reside in your household?

```
with(data, summary(Q33))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    1.00    2.00    2.00    2.58    3.00   14.00      28
```

```
# by ethnicity
with(data, by(Q33, PPETHM, summary))
```

```
## PPETHM: 2+ Races, Non-Hispanic
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   2.000   2.709   3.000   7.000       1
## ------------------------------------------------------------
## PPETHM: Black, Non-Hispanic
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   1.000   2.000   2.544   3.000  13.000       2
## ------------------------------------------------------------
## PPETHM: Hispanic
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   3.000   2.903   4.000   9.000       6
## ------------------------------------------------------------
## PPETHM: Other, Non-Hispanic
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   3.000   2.946   4.000   7.000       1
## ------------------------------------------------------------
## PPETHM: White, Non-Hispanic
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   2.000   2.000   2.509   3.000  14.000      18
```

Household Members

HHM1

Q35. What is the gender of this member of the household? Remember, this relates to HHM1_Name who is HHM1_AGE years old.

```
with(data, table(Q35))
```

```
## Q35
## Female   Male
##    799    859
```

Q36. On average, how many days per week does this member of your household work or attend day care or school outside of your home?

```
with(data, summary(Q36))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   4.000   2.874   5.000   7.000     571
```

Q37. On average, how many days per week does this member of your household participate in social activities outside of your home?

```
with(data, summary(Q37))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   2.000   2.098   3.000   7.000     663
```

Q38. On average, how many days per week does this member of your household use public transportation?

```
with(data, summary(Q38))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.0000  0.0000  0.3909  0.0000  7.0000     582
```

Q39. How frequently does this member of your household visit a doctor's office for wellness appointments?

```
with(data, summary(Q39))
```

```
##              Don_t know Less than once per year More than once per year
##                      84                     222                     593
##                   Never           Once per year                   NA's
##                     100                     656                     513
```

Q40. How frequently does this member of the household get sick in a typical year?

```
with(data, summary(Q40))
```

```
##  1 to 2 times  3 to 5 times 6 to 10 times    Don_t know  More than 10
##          1025           271            32            87            13
##         Never          NA's
##           226           514
```

Q41. How many times has this member of your household had influenza or another respiratory illness in the last two years?

```
with(data, summary(Q41))
```

```
##     2 times     3 times  Don_t know More than 3       Never        Once
##         191          60         158          39         807         400
##        NA's
##         513
```

Q42. Does this member of your household get an annual influenza vaccine?

```
with(data, summary(Q42))
```

```
##     Don_t know      No, never    Yes, always Yes, sometimes           NA's
##            166            567            661            263            511
```

HHM2

Q43. What is the gender of this member of the household? Remember, this relates to HHM1_Name who is HHM1_AGE years old.

```
with(data, summary(Q43))
```

```
## Female   Male   NA's
##    388    431   1349
```

Q44. On average, how many days per week does this member of your household work or attend day care or school outside of your home?

```
with(data, summary(Q44))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   5.000   3.669   5.000   7.000    1383
```

Q45. On average, how many days per week does this member of your household participate in social activities outside of your home?

```
with(data, summary(Q45))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   1.000   2.000   2.395   4.000   7.000    1419
```

Q46. On average, how many days per week does this member of your household use public transportation?

```
with(data, summary(Q46))
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.0000  0.0000  0.5727  0.0000  7.0000    1391
```

Q47. How frequently does this member of your household visit a doctor's office for wellness appointments?

```
with(data, summary(Q47))
```

```
##              Don_t know Less than once per year More than once per year
##                      55                     125                     232
##                   Never           Once per year                   NA's
##                      59                     353                    1344
```

Q48. How frequently does this member of the household get sick in a typical year?

```
with(data, summary(Q48))
```

```
##  1 to 2 times  3 to 5 times 6 to 10 times    Don_t know  More than 10
##           490           153            25            66             7
##         Never          NA's
##            83          1344
```

Q49. How many times has this member of your household had influenza or another respiratory illness in the last two years?

```
with(data, summary(Q49))
```

```
##     2 times     3 times  Don_t know More than 3       Never        Once
##          91          32          93          21         403         183
##        NA's
##        1345
```

Q50. Does this member of your household get an annual influenza vaccine?

```
with(data, summary(Q50))
```

```
##     Don_t know      No, never    Yes, always Yes, sometimes           NA's
##            100            317            275            132           1344
```