plan_of_analysis

Arielle Herman

2/11/2022

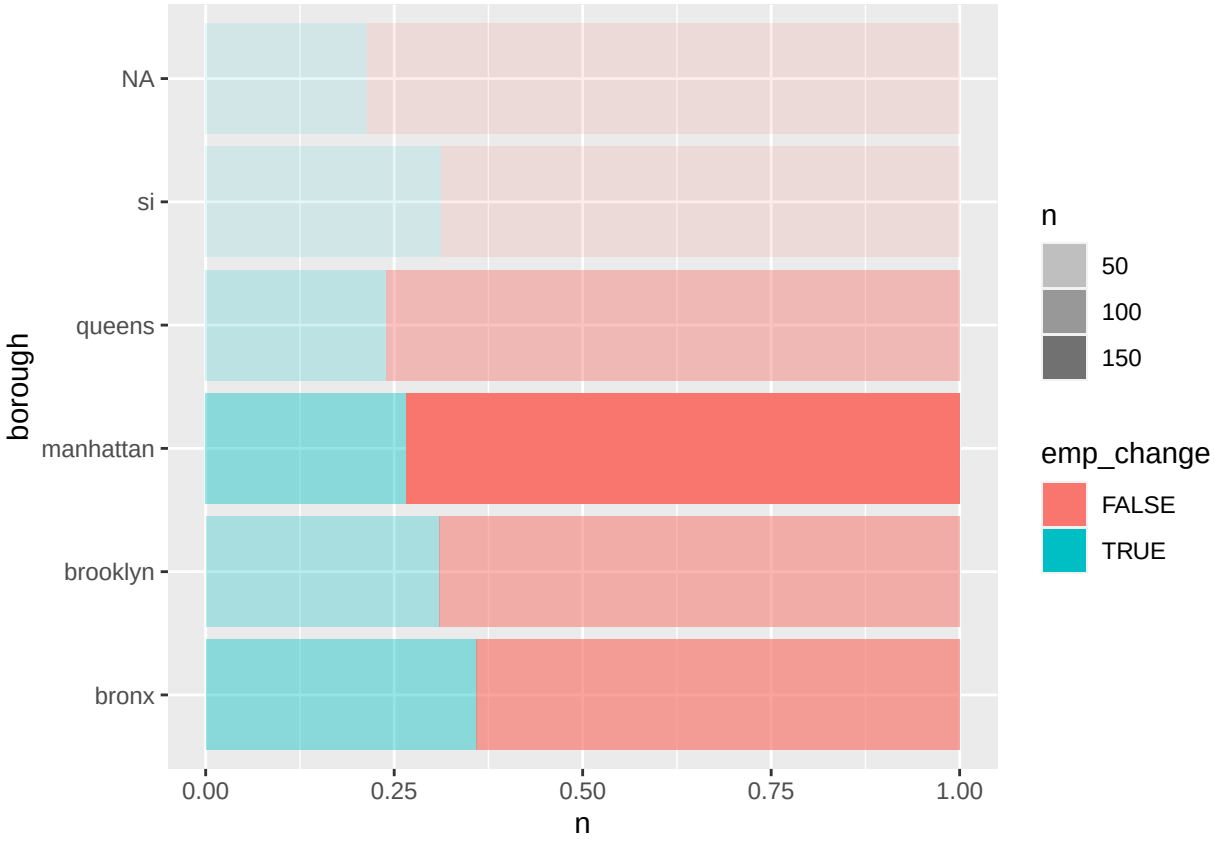
```
library(tidyverse)
library(haven)
wrangled <- read_dta("../data/output/wrangled20220211.dta") %>%
  mutate_if(is.character, na_if, "")
```

employment

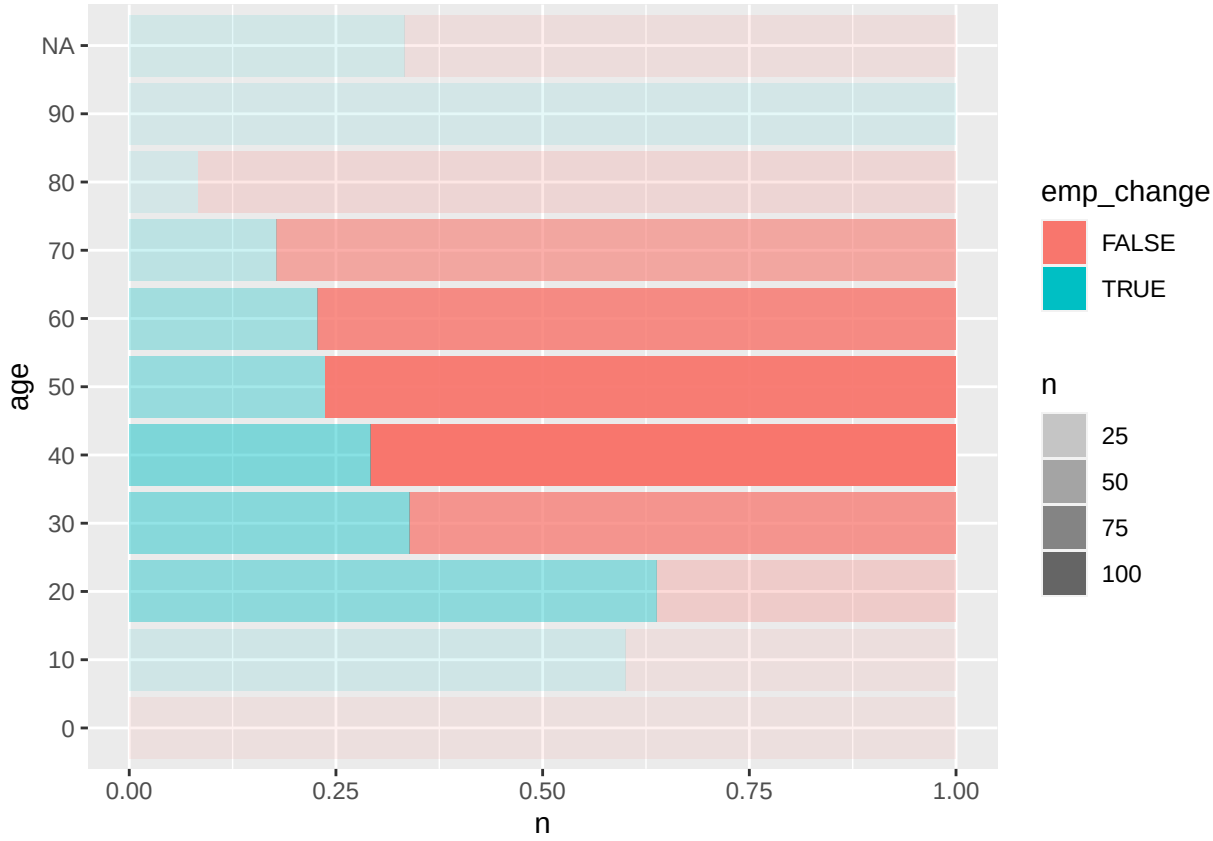
```
demographics <- c("borough", "age", "gen", "not_eng", "mar", "sch_level",
                  "hh_64_bi", "hh_ch_0_17", "inc_dist", "res_type") # k

lapply(demographics, function(dem) {
  sym_dem <- sym(dem)
  reshape <- wrangled %>% count(emp_change = emp_b != emp_a, !!sym_dem)
  if(is.character(reshape[[dem]])) {
    reshaped <- reshape %>% arrange(n)
  } else if(haven::is.labelled(reshape[[dem]])){
    reshaped <- reshape %>% arrange(!!sym_dem) %>%
      mutate_at(vars(!!sym_dem), haven::as_factor)
  } else if (dem == "age") {
    reshaped <- reshape %>% mutate_at(vars(!!sym_dem), ~as.factor(floor(./10)*10)) %>%
      group_by(emp_change, !!sym_dem) %>% summarize(n = sum(n))
  } else {
    reshaped <- reshape %>% mutate_at(vars(!!sym_dem), ~as.factor(.)) %>% arrange(n)
  }
  p <- reshaped %>%
    ggplot(aes(x = n, y = !!sym_dem, fill = emp_change, alpha = n)) +
    geom_col(position = "fill")# + xlab(dem)
  return(p)
})
```

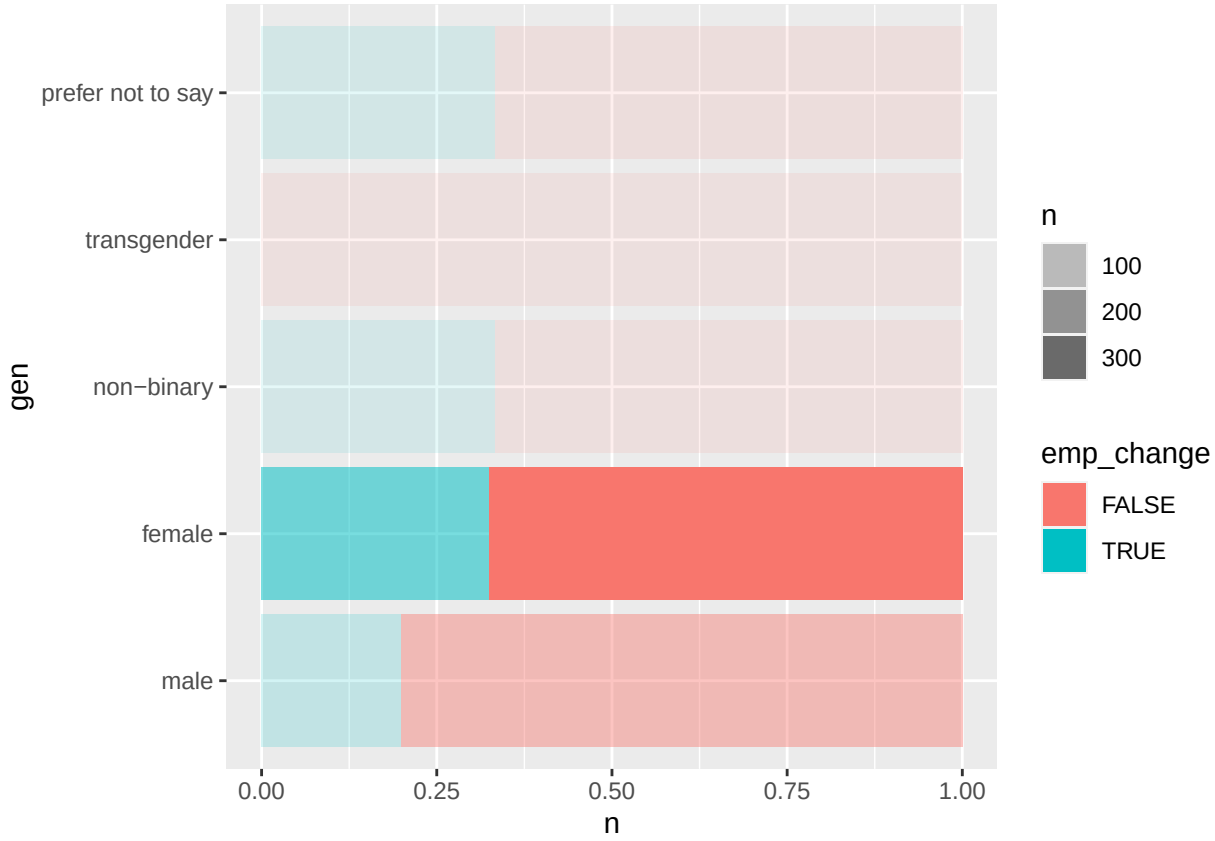
```
## [[1]]
```


NA
si
queens
manhattan
brooklyn
bronx
borough
0.00
0.25
0.50
0.75
1.00
n
n
50
100
150
emp_change
FALSE
TRUE

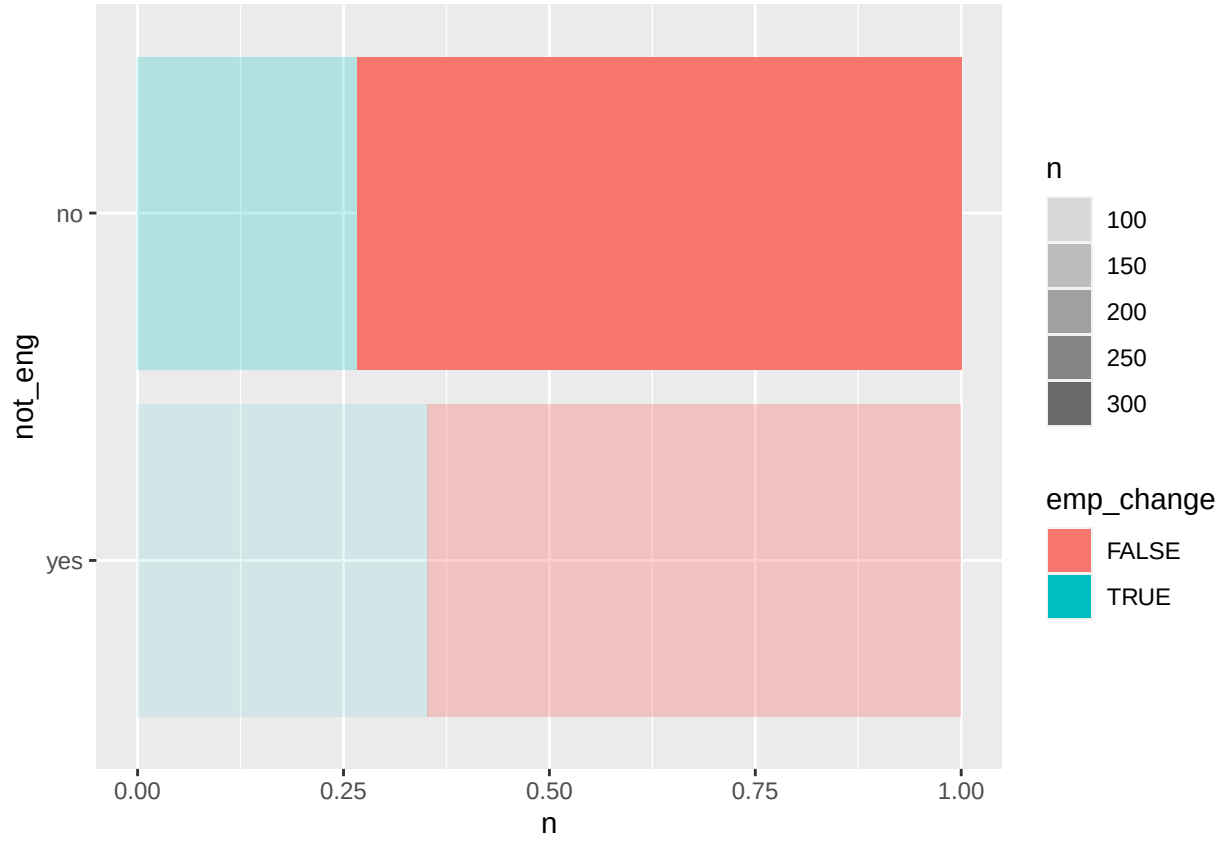
```
##
## [[2]]
```


NA
90
80
70
60
50
40
30
20
10
0
age
0.00
0.25
0.50
0.75
1.00
n
emp_change
FALSE
TRUE
n
25
50
75
100

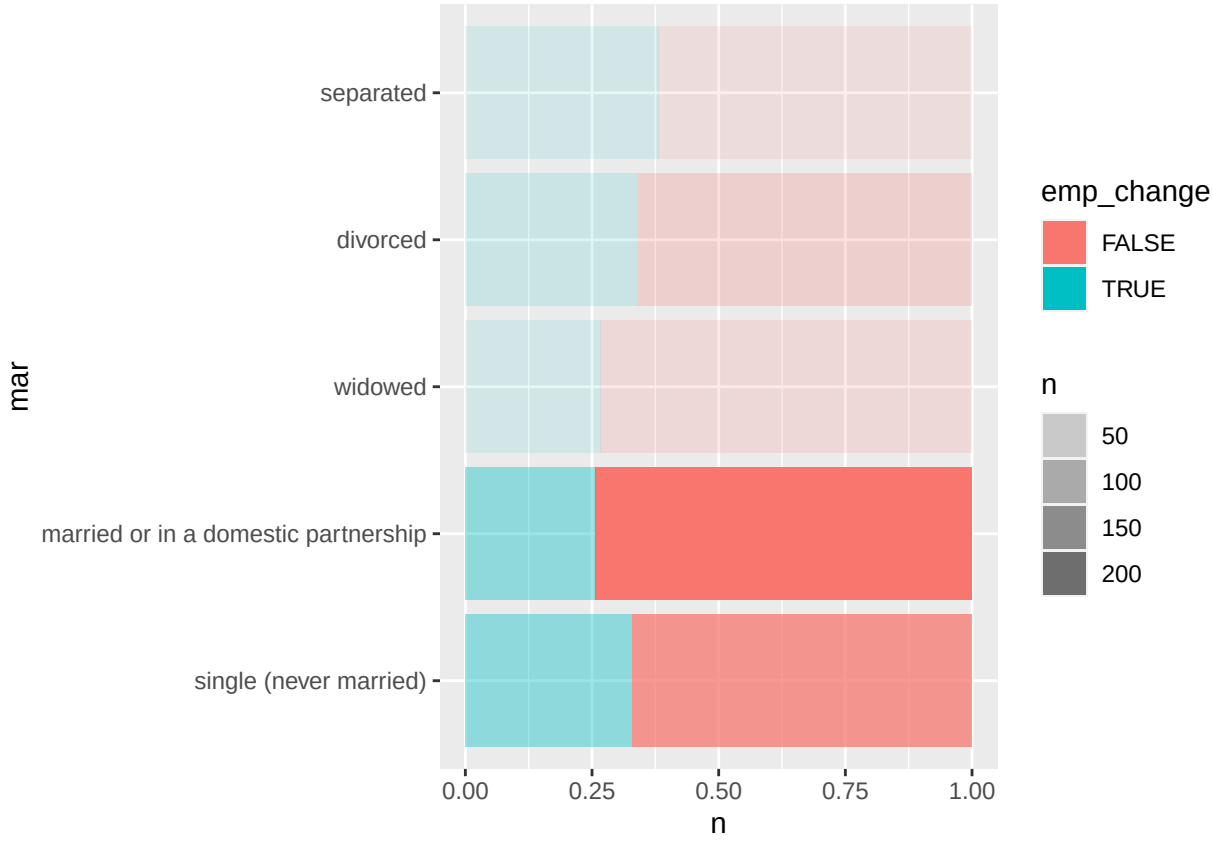
```
##
## [[3]]
```


prefer not to say
transgender
non−binary
female
male
gen
0.00
0.25
0.50
0.75
1.00
n
n
100
200
300
emp_change
FALSE
TRUE

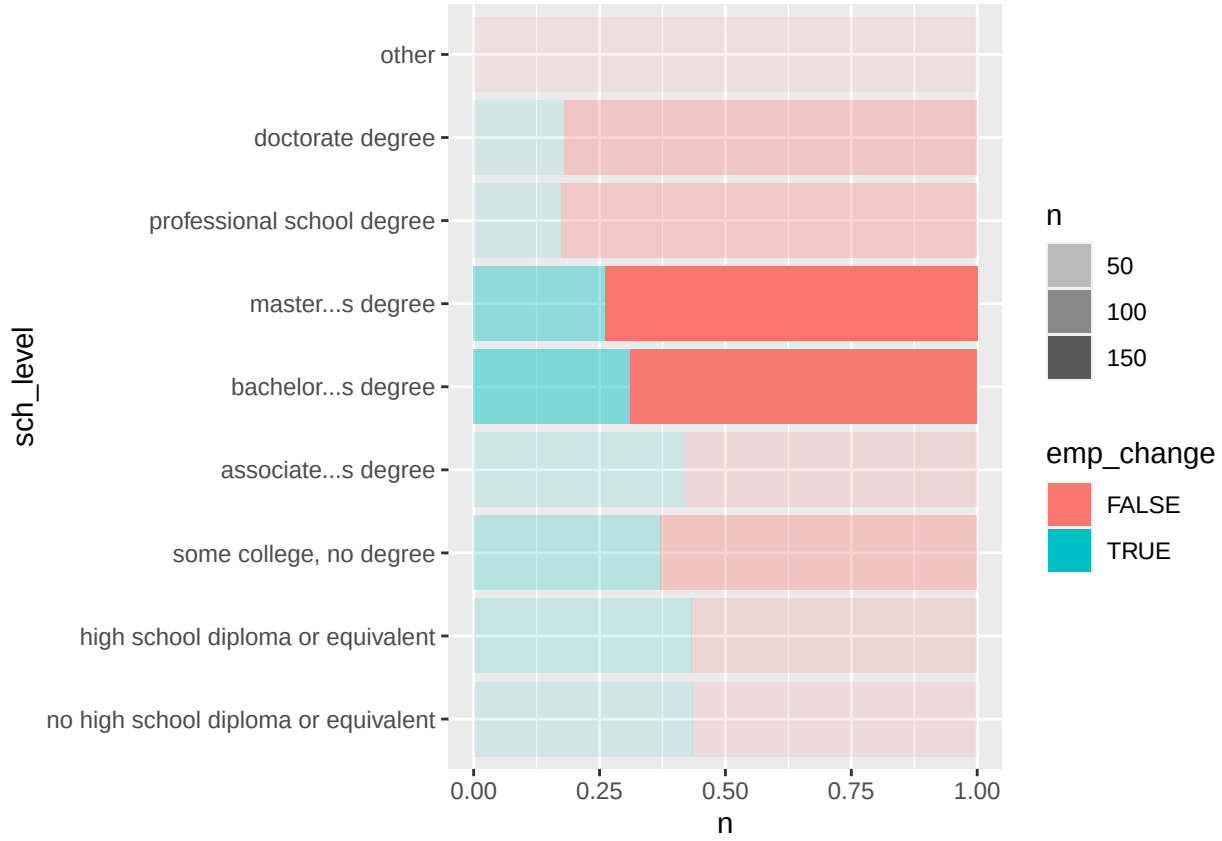
```
##
## [[4]]
```


not_eng
no
yes
0.00
0.25
0.50
0.75
1.00
n
n
100
150
200
250
300
emp_change
FALSE
TRUE

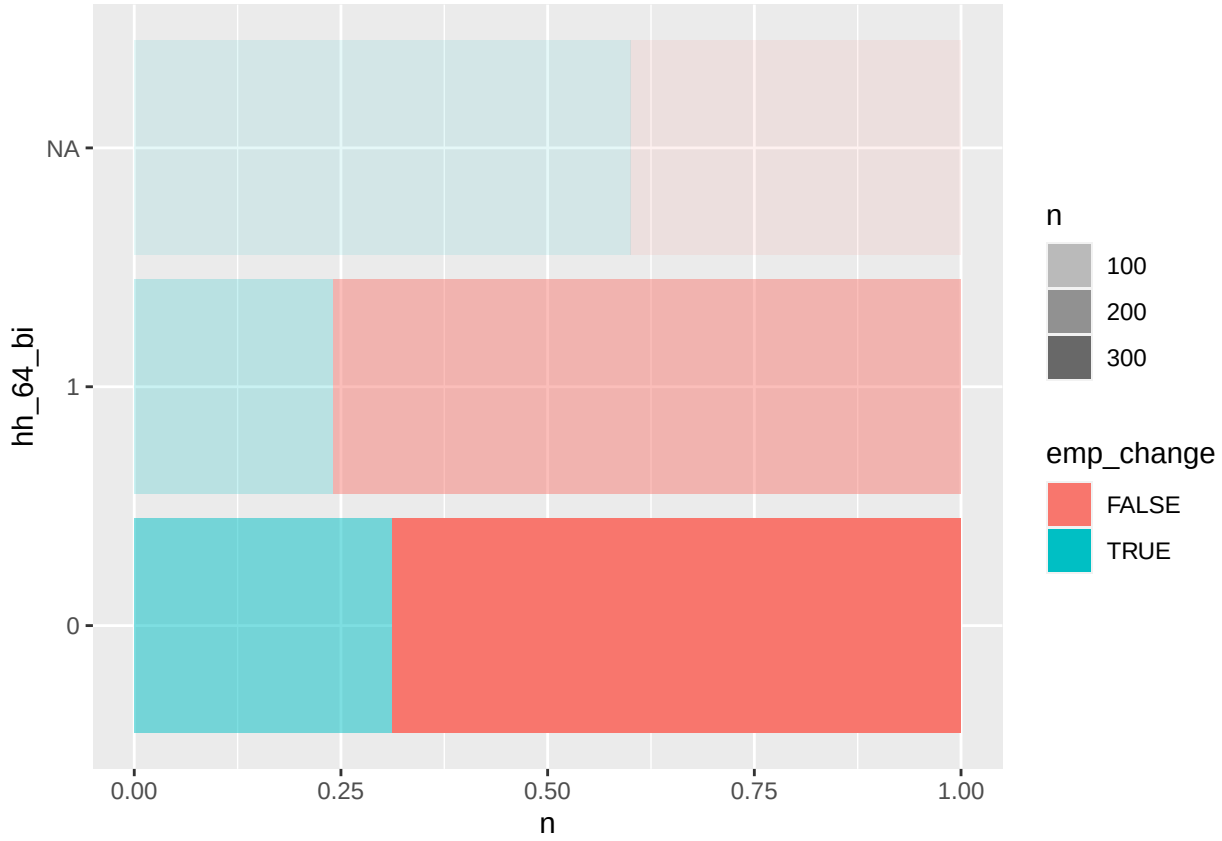
```
##
## [[5]]
```


separated
divorced
widowed
married or in a domestic partnership
single (never married)
mar
0.00
0.25
0.50
0.75
1.00
n
emp_change
FALSE
TRUE
n
50
100
150
200

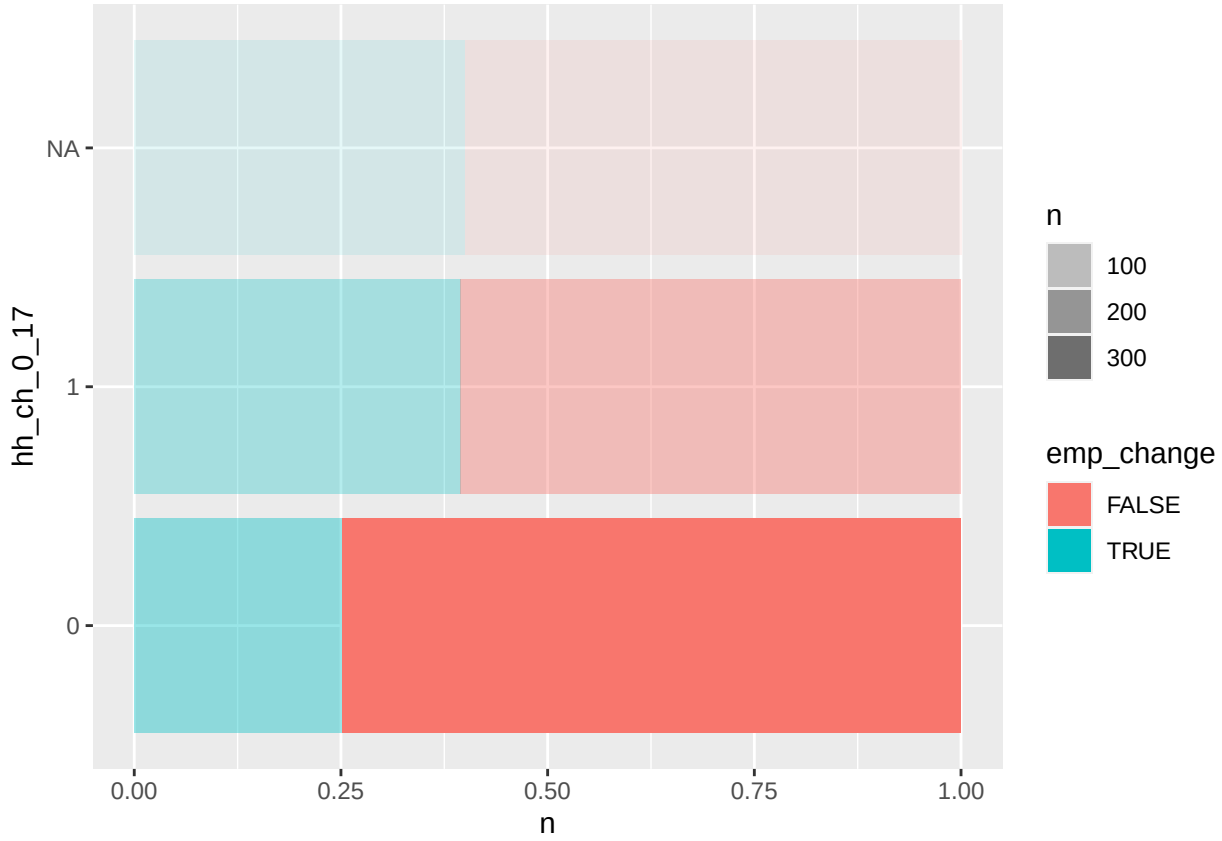
```
##
## [[6]]
```


other
doctorate degree
professional school degree
master...s degree
bachelor...s degree
associate...s degree
some college, no degree
high school diploma or equivalent
no high school diploma or equivalent
sch_level
0.00
0.25
0.50
0.75
1.00
n
n
50
100
150
emp_change
FALSE
TRUE

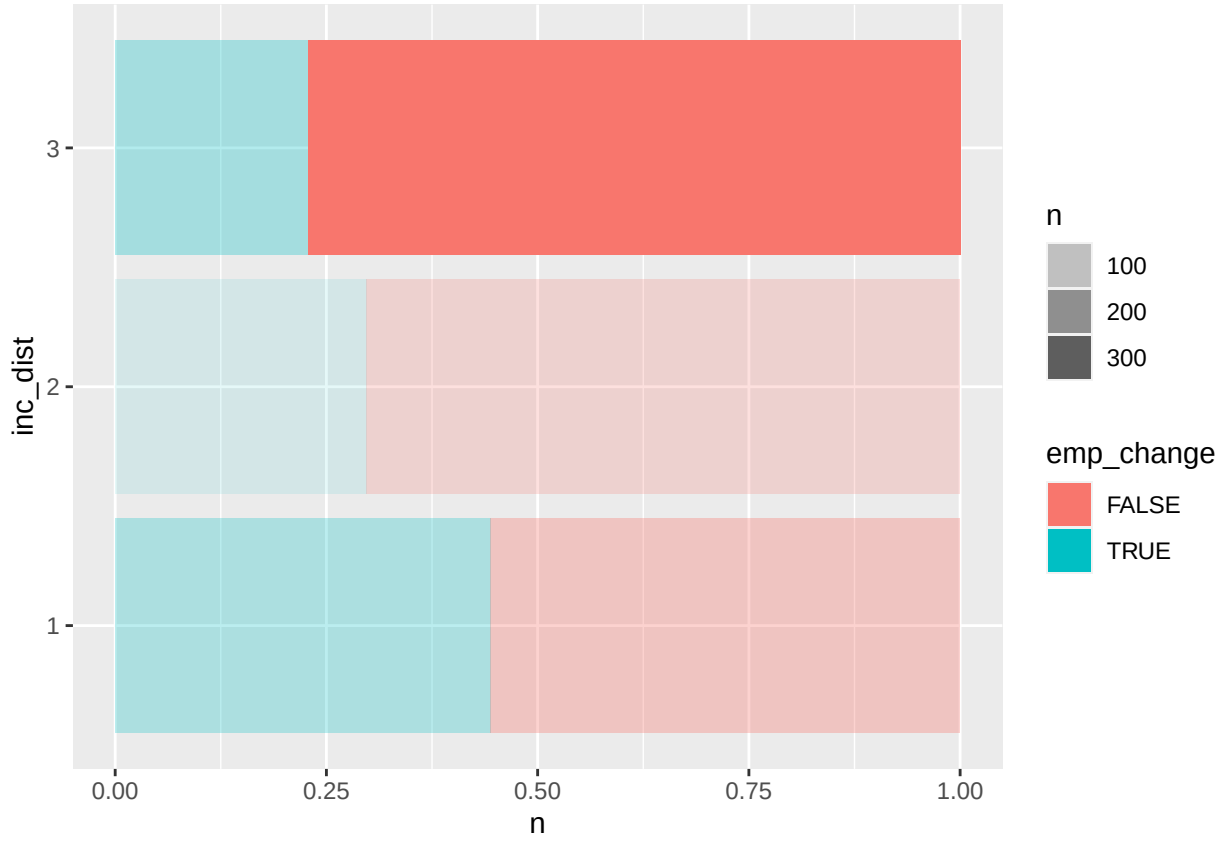
```
##
## [[7]]
```


NA
1
0
hh_64_bi
0.00
0.25
0.50
0.75
1.00
n
n
100
200
300
emp_change
FALSE
TRUE

```
##
## [[8]]
```

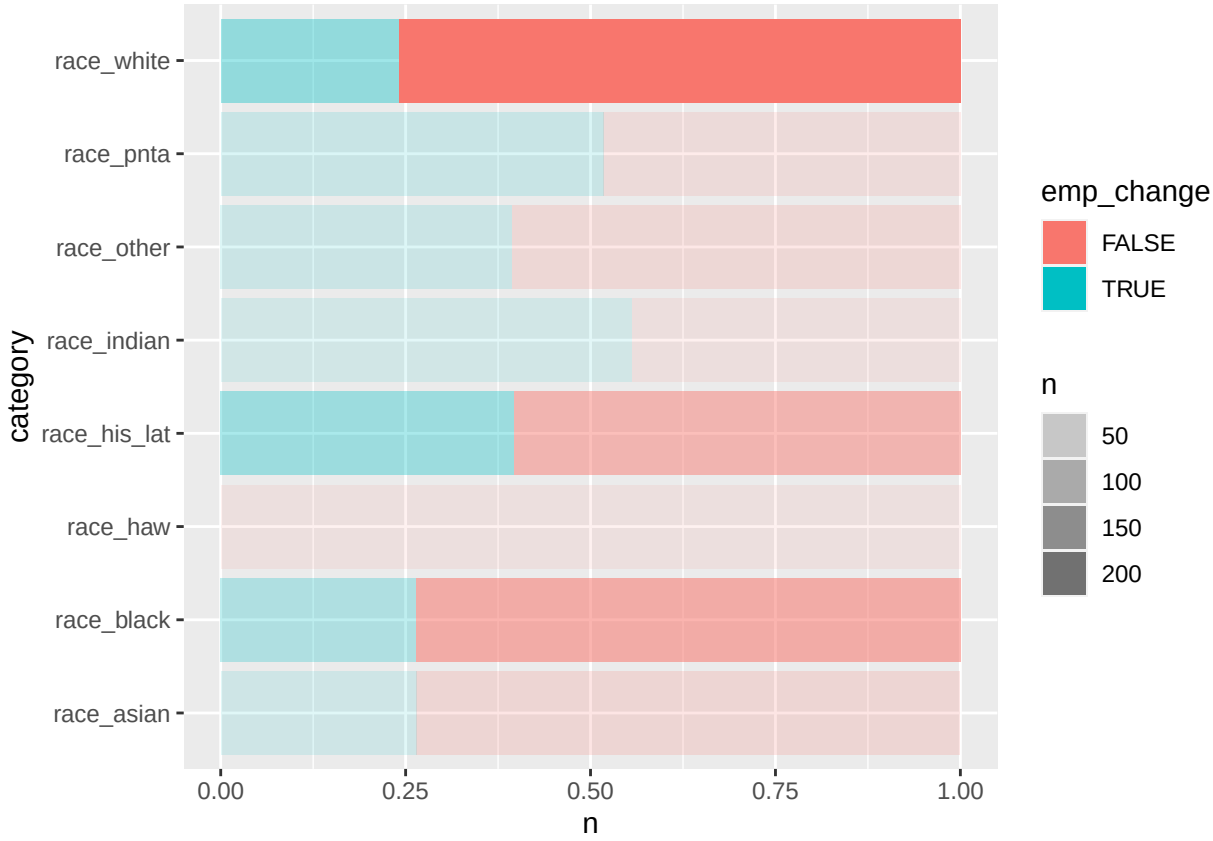

NA
1
0
hh_ch_0_17
0.00
0.25
0.50
0.75
1.00
n
n
100
200
300
emp_change
FALSE
TRUE

```
##
## [[9]]
```


inc_dist
3
2
1
0.00
0.25
0.50
0.75
1.00
n
n
100
200
300
emp_change
FALSE
TRUE

```
##
## [[10]]
```


```
# does not yet fix for double counting
wrangled %>%
  group_by(responseid, emp_change = emp_b != emp_a) %>%
  select(responseid, contains("race_"), -race_text, -race_twomore) %>%
  pivot_longer(cols = contains("race_"), names_to = "category", values_to = "value") %>%
  filter(value != 0) %>% ungroup %>% count(category, emp_change) %>%
  ggplot(aes(x = n, y = category, fill = emp_change, alpha = n)) + geom_col(position = "fill")
```


race_white
race_pnta
race_other
race_indian
race_his_lat
race_haw
race_black
race_asian
category
0.00
0.25
0.50
0.75
1.00
n
emp_change
FALSE
TRUE
n
50
100
150
200