Stat 245 – Michigan Housing

Sharon Velpula

December 09, 2022

Read in Data

```
hdma_mi_20 <- read_csv('https://sldr.netlify.app/data/hdma-mi-20.csv', show_col_types = FALSE)

head(hdma_mi_20)
```

```
## # A tibble: 6 x 99
##   activity_year lei         `derived_msa-md` state_code county_code census_tract
##           <dbl> <chr>                  <dbl> <chr>            <dbl>        <dbl>
## 1          2020 549300YIQ7~            47664 MI               26099  26099210000
## 2          2020 549300YIQ7~            47664 MI               26125  26125138100
## 3          2020 549300YIQ7~            19804 MI               26163  26163564502
## 4          2020 549300YIQ7~            19804 MI               26163  26163542200
## 5          2020 549300YIQ7~            99999 MI               26009  26009960300
## 6          2020 549300YIQ7~            99999 MI               26151  26151971000
## # ... with 93 more variables: conforming_loan_limit <chr>,
## #   derived_loan_product_type <chr>, derived_dwelling_category <chr>,
## #   derived_ethnicity <chr>, derived_race <chr>, derived_sex <chr>,
## #   action_taken <dbl>, purchaser_type <dbl>, preapproval <dbl>,
## #   loan_type <dbl>, loan_purpose <dbl>, lien_status <dbl>,
## #   reverse_mortgage <dbl>, `open-end_line_of_credit` <dbl>,
## #   business_or_commercial_purpose <dbl>, loan_amount <dbl>, ...
```

Data Wrangling

```
hdma_mi_20 <- hdma_mi_20 |>
  mutate(action_taken = case_when(action_taken == 1 ~ 'Loan granted', action_taken == 3 ~ 'Loan denied'

mosaic::tally(~action_taken, data = hdma_mi_20)
```

```
## action_taken
##  Loan denied Loan granted
##         8338       100333
```

```
hdma_mi_20_sex <- hdma_mi_20 %>%
  filter(derived_sex %in% c("Male", "Female", "Joint", "Sex Not Available"))
```

Calculate average denial rate for every sex

```
hdma_mi_20_denied <- hdma_mi_20 |>
  group_by(derived_sex) |>
  mutate(rate_of_denial=paste0(round(count(action_taken=='Loan denied')/sum(action_taken=='Loan granted

select("derived_sex", "rate_of_denial")

glimpse(hdma_mi_20_denied)
```

```
## Rows: 108,671
## Columns: 2
## Groups: derived_sex [4]
## $ derived_sex    <chr> "Male", "Joint", "Female", "Sex Not Available", "Sex No~
## $ rate_of_denial <chr> "8", "5", "9", "13", "13", "8", "9", "9", "8", "9", "13~
```

```
unique(hdma_mi_20_denied$rate_of_denial)
```

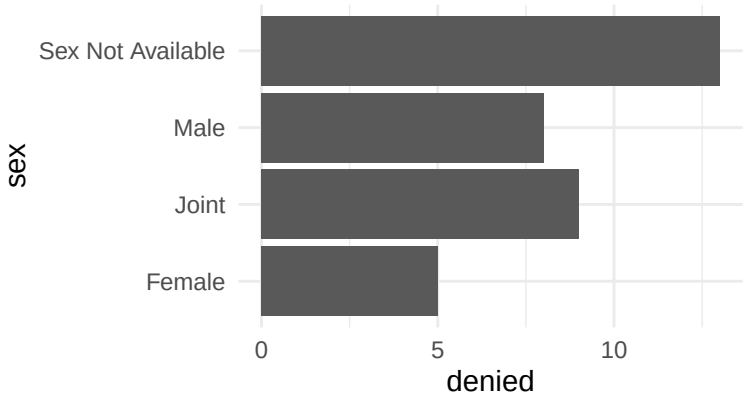
```
## [1] "8"  "5"  "9"  "13"
```

New Dataframe including only sex and denied numbers

```
data <- data.frame(
  sex = c("Male", "Female", "Joint", "Sex Not Available"),
  denied = c(8, 5, 9, 13)
)
```

Graph

```
data %>%
  ggplot() +
  aes(y= sex, x = denied) +
  geom_col(position = "dodge") +
  theme_minimal()
```



I attempted to recreate the line graph that shows the rate of denial in response to credit score by The Markup. But while going through the process, I wanted to show the rate of denial in response to credit score by sex. But I was not able to do all the calculations properly so my end result is basically a bar plot showing denied

numbers by sex. Looks like those whose sex was not available were denied the most while females were denied the least.