SPSS Assignment 1

JP

8/16/2021

What Each Assignment Will Look Like

Create a `null` and `alternative` hypothesis

Rename `gender` into `sex`

Conduct descriptive statistics for the following variables

- also include the appropriate descriptive visualization for each variable below
- `sex`
- `mtuas internet q1`
- `mtuas internet q2`
- `mtuas internet q3`
- `mtuas internet q4`
- `avg internet score variable`

Conduct inferential statistic (independent samples t-test)

- screenshot the table with the independent samples t-test

Tell me whether or not your finding is statistically significant

find the t-critical value for a two-tailed test

Show some type of visualization that shows the difference in the two groups on internet use

- examples of what it should look like are below

Reject/Retain the Null Hypothesis?

Write up on your findings

```
jp <- jp %>%
  mutate(sex = factor(ccc_gender),
         internet_avg = (mtuas_internet_q1 + mtuas_internet_q2 + mtuas_internet_q3 + mtuas_interne
```

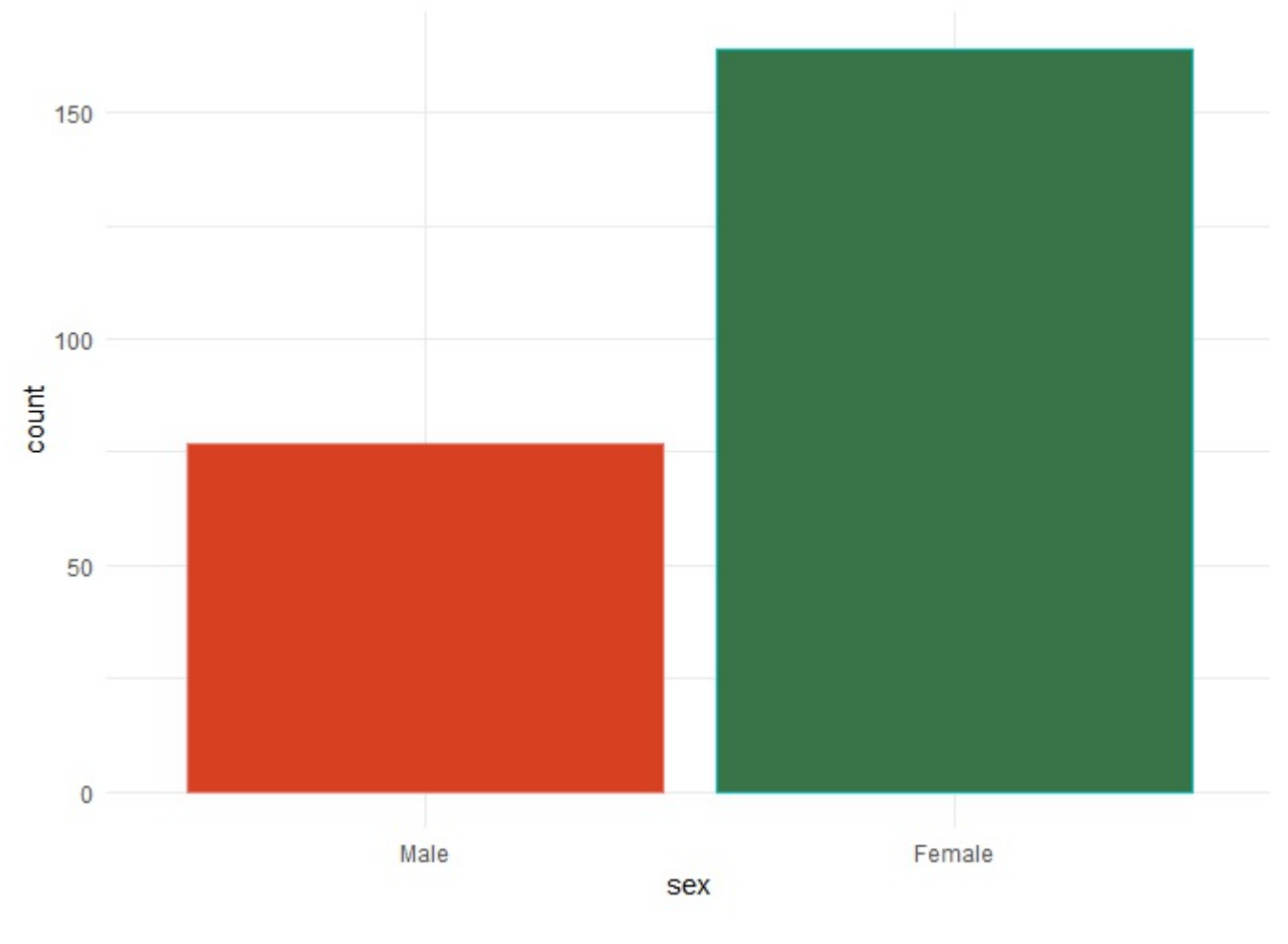
```
psych::describeBy(jp$internet_avg, jp$sex, na.rm = TRUE)
```

```
##
##  Descriptive statistics by group
## group: 1
##    vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 77 5.95 1.95   5.75     5.9 1.85   2  10     8 0.27    -0.53 0.22
## ---------------------------------------------------------------------------
## group: 2
##    vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 164 6.61 2.17      6    6.58 2.22 2.5  10   7.5 0.26     -1.2 0.17
```

```
psych::describe(jp$internet_avg, na.rm = TRUE)
```

```
##    vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 241  6.4 2.12   5.75    6.35 1.85   2  10     8 0.29    -0.96 0.14
```

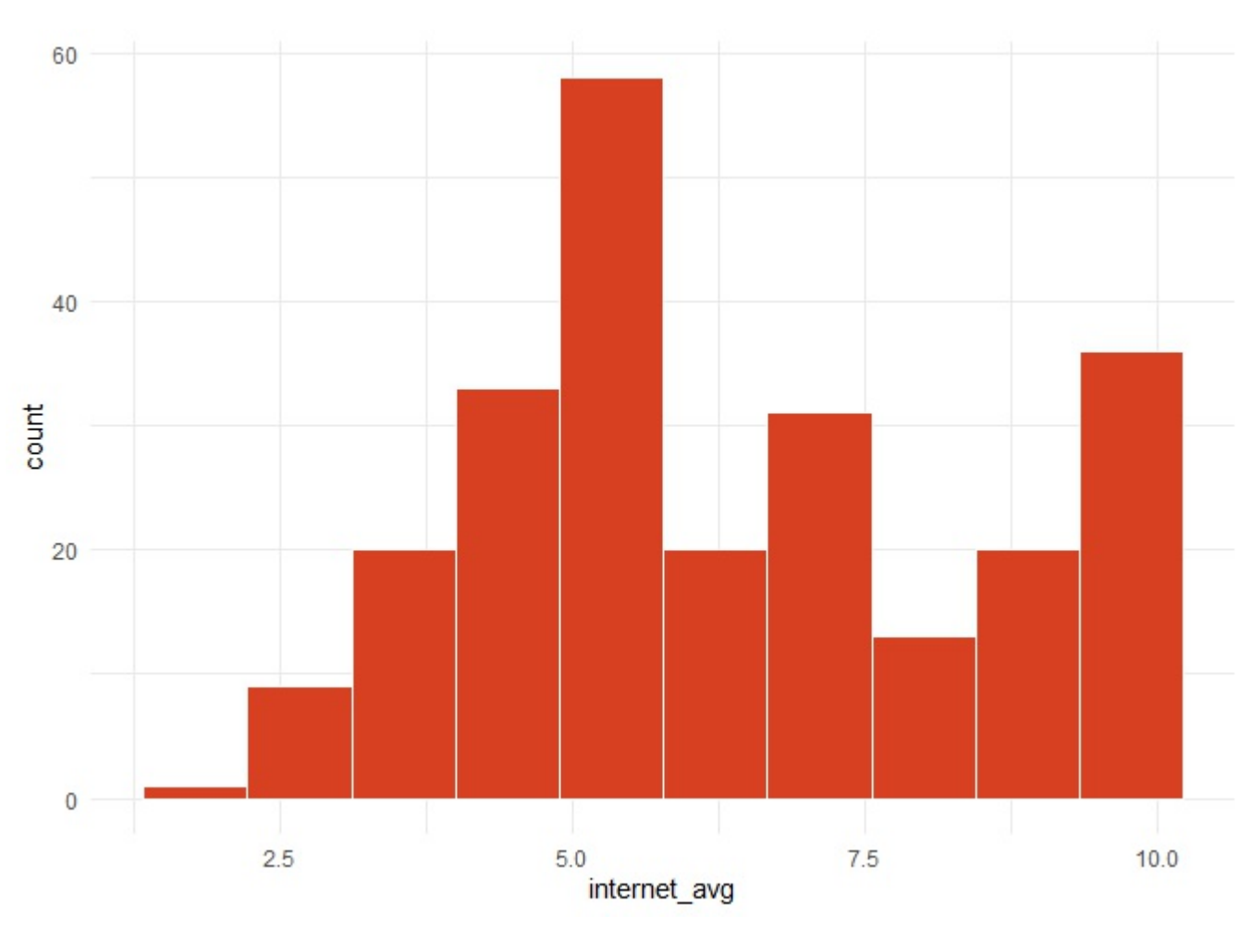
```
jp %>%
  mutate(sex = recode(sex, "1" = "Male",
                      "2" = "Female")) %>%
  ggplot(aes(sex)) +
  geom_bar(aes(color = sex, fill = sex)) +
  scale_fill_manual(values = c("#d74122", "#387448")) +
  theme(legend.position = "none")
```




```
jp %>%
  group_by(sex) %>%
  count()
```

```
## # A tibble: 2 x 2
## # Groups:   sex [2]
##   sex       n
##   <fct> <int>
## 1 1        77
## 2 2       164
```

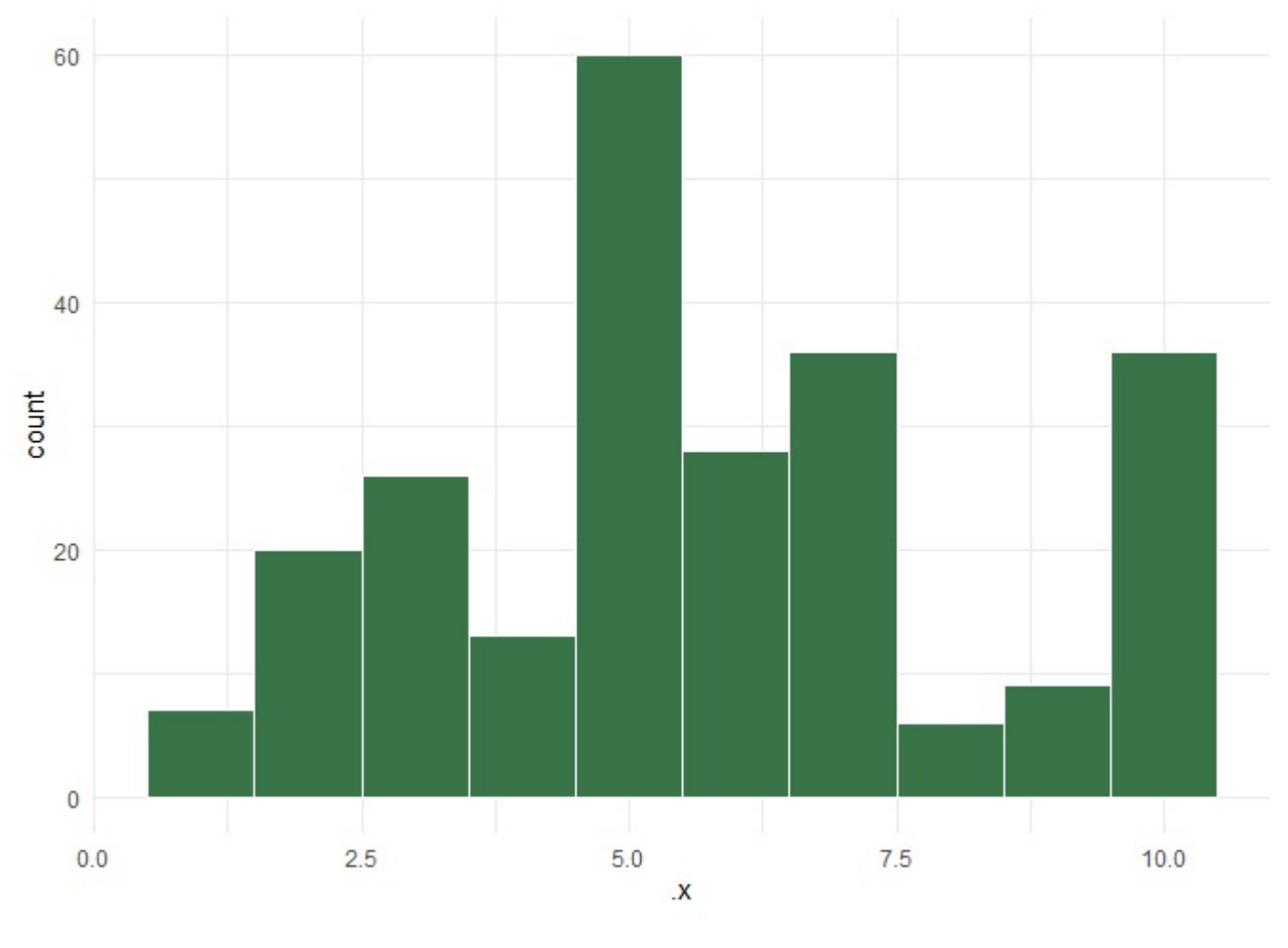
```
jp %>%
  ggplot(aes(internet_avg)) +
  geom_histogram(color = "white", fill = "#d74122", bins = 10)
```



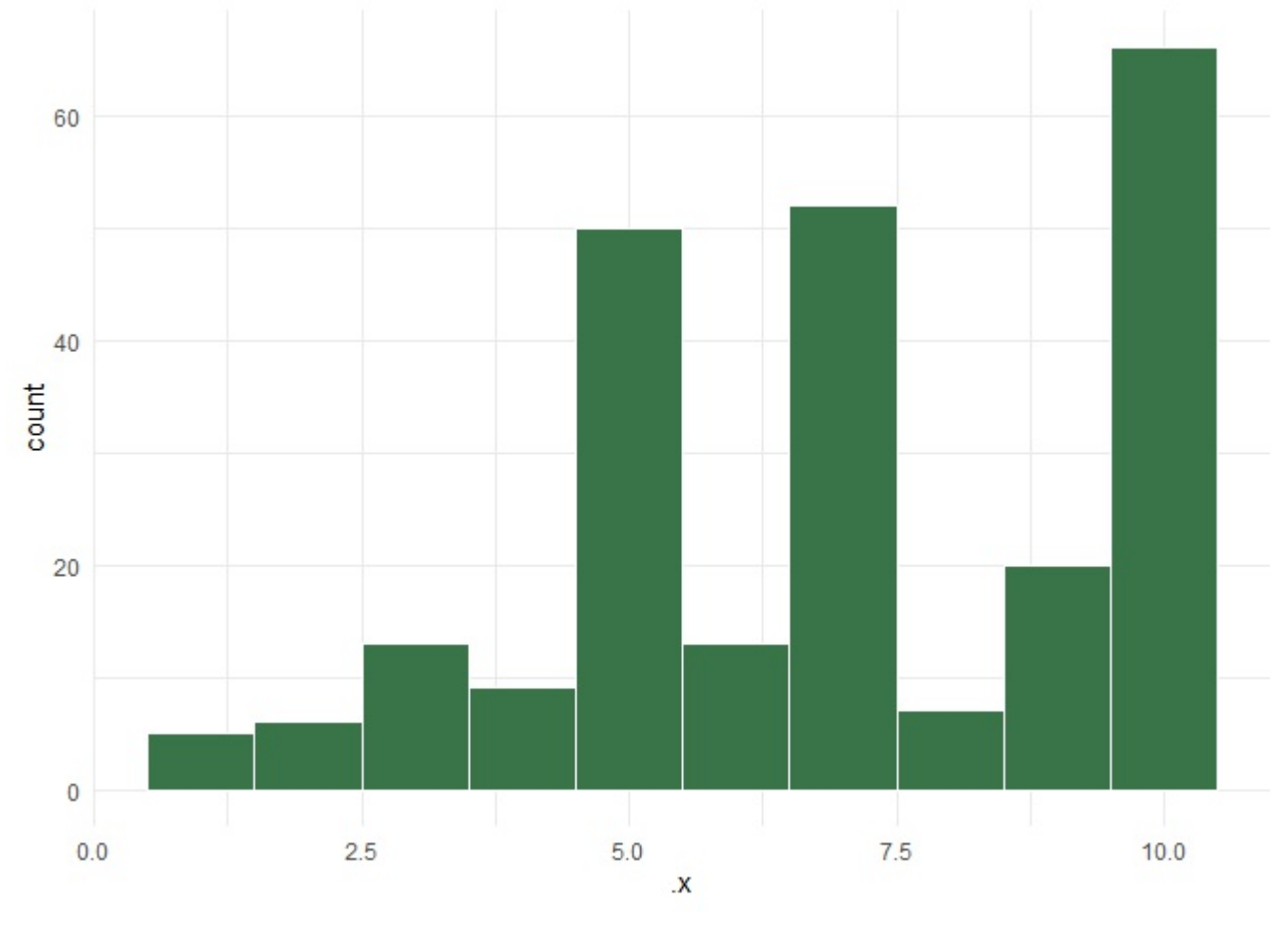

```
internet_only <- jp %>%
  dplyr::select(mtuas_internet_q1:mtuas_internet_q4)

map(internet_only, ~ggplot(data = internet_only, aes(.x)) + geom_histogram(color = "white", fill =
```

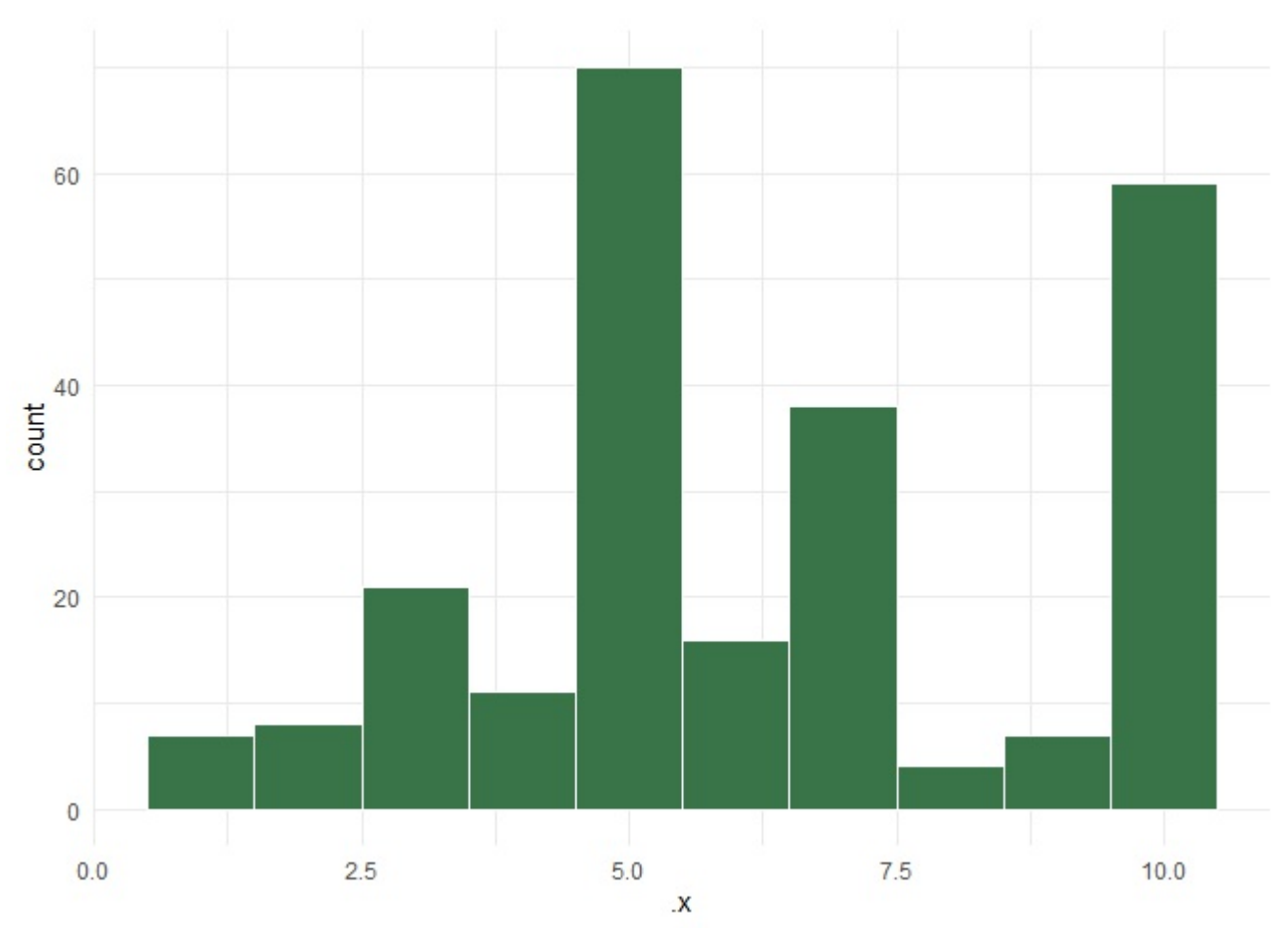
```
## $mtuas_internet_q1
```



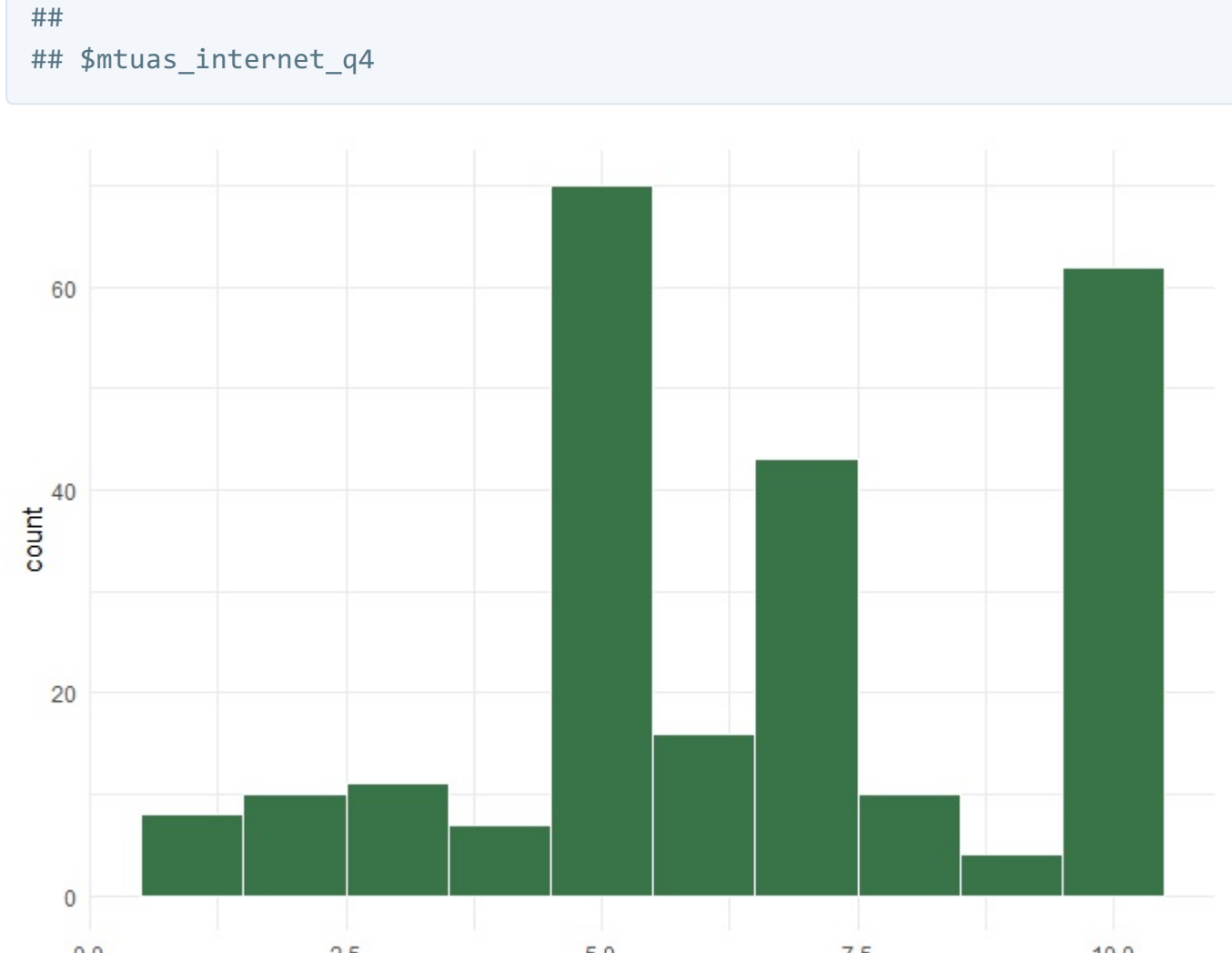

```
##
## $mtuas_internet_q2
```




```
##
## $mtuas_internet_q3
```




```
##
## $mtuas_internet_q4
```




```
car::leveneTest(internet_avg ~ sex, data = jp)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value  Pr(>F)
## group   1  3.6245 0.05814 .
##       239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

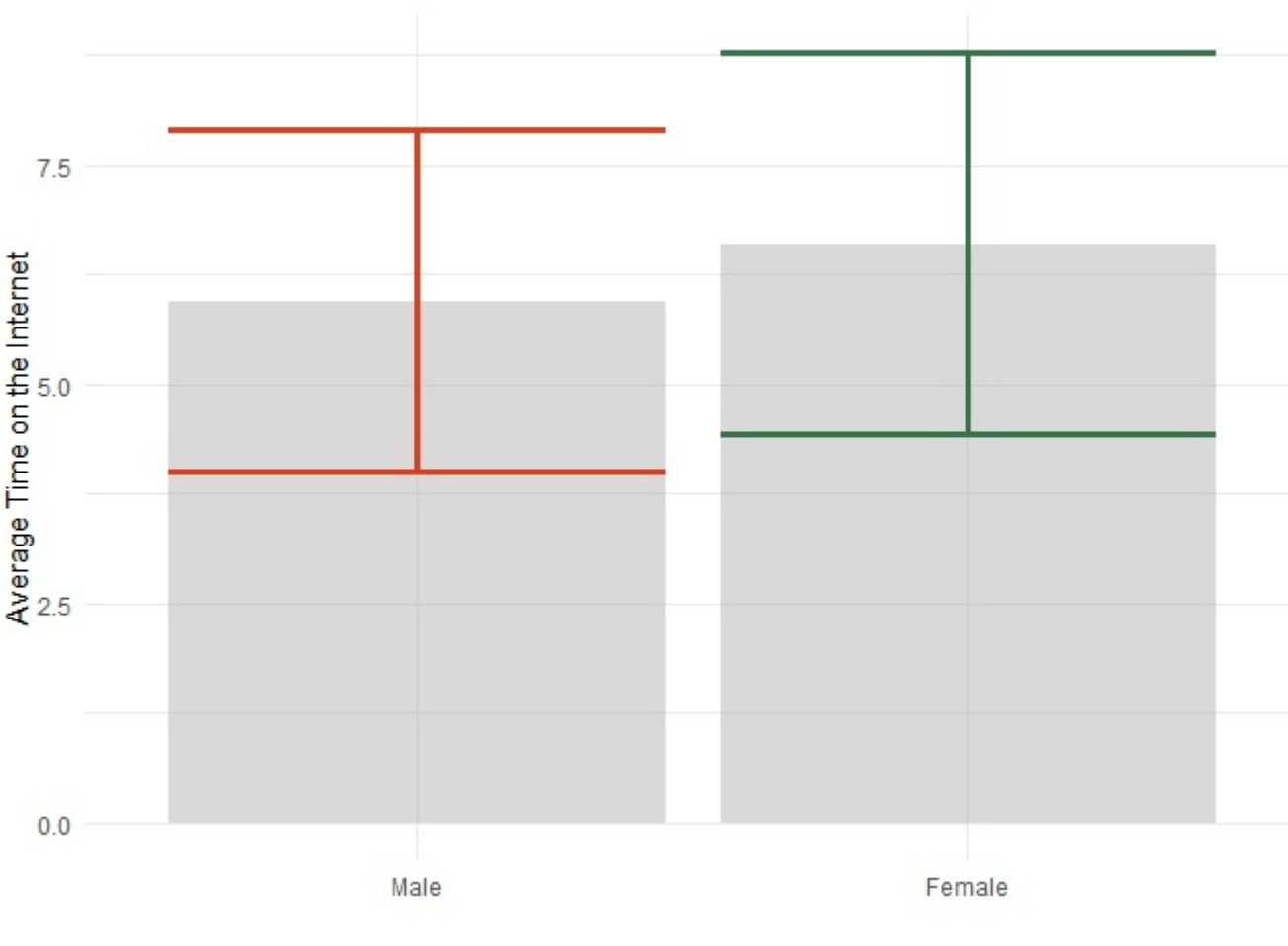
```
t_test_finding <- t.test(internet_avg ~ sex, data = jp, var.equal = TRUE)
t_test_finding
```

```
##
##  Two Sample t-test
##
## data:  internet_avg by sex
## t = -2.2616, df = 239, p-value = 0.02462
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.22952129 -0.08474066
## sample estimates:
## mean in group 1 mean in group 2
##        5.948052        6.605183
```

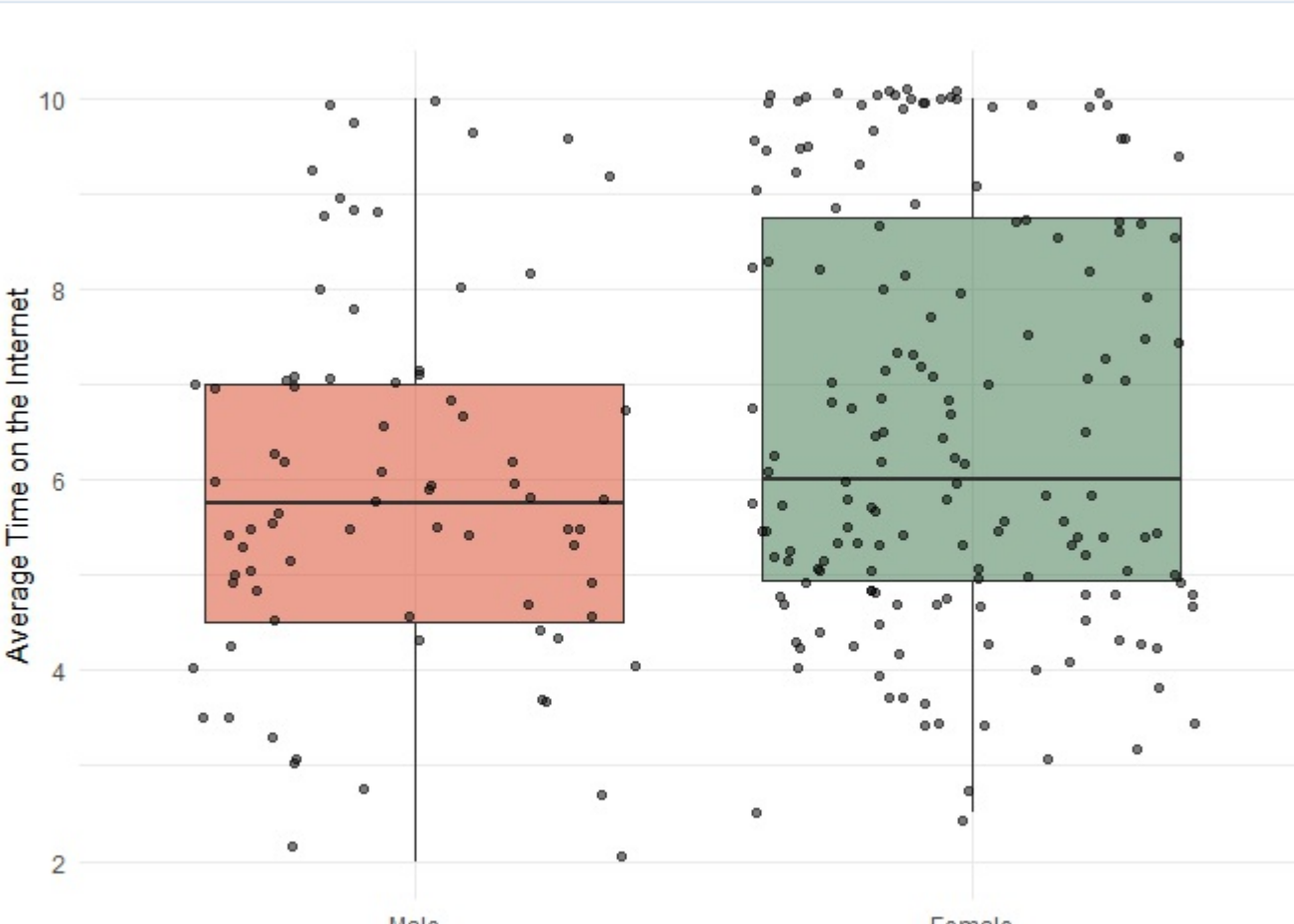
```
t.test(internet_avg ~ sex, data = jp)
```

```
##
##  Welch Two Sample t-test
##
## data:  internet_avg by sex
## t = -2.3502, df = 164.11, p-value = 0.01995
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.2092251 -0.1050368
## sample estimates:
## mean in group 1 mean in group 2
##        5.948052        6.605183
```

```
jp %>%
  mutate(sex = recode(sex, "1" = "Male",
                      "2" = "Female")) %>%
  group_by(sex) %>%
  summarize(mean_internet = mean(internet_avg),
            sd_internet = sd(internet_avg)) %>%
  ungroup() %>%
  ggplot(aes(sex, mean_internet)) +
  geom_col(fill = "gray70", alpha = .5) +
  geom_errorbar(aes(ymin = mean_internet - sd_internet,
                    ymax = mean_internet + sd_internet,
                    color = sex), size = 1.25) +
  scale_color_manual(values = c("#d74122", "#387448")) +
  theme(legend.position = "none") +
  labs(x = "",
       y = "Average Time on the Internet")
```




```
jp %>%
  mutate(sex = recode(sex, "1" = "Male",
                      "2" = "Female")) %>%
  ggplot(aes(sex, internet_avg)) +
  geom_boxplot(aes(fill = sex), alpha = .5) +
  geom_jitter(color = "black", alpha = .5) +
  scale_fill_manual(values = c("#d74122", "#387448")) +
  theme(legend.position = "none") +
  labs(x = "",
       y = "Average Time on the Internet")
```




```
broom::tidy(t_test_finding)
```

```
## # A tibble: 1 x 10
##   estimate estimate1 estimate2 statistic p.value parameter conf.low conf.high method        alt
##      <dbl>     <dbl>     <dbl>     <dbl>   <dbl>     <dbl>    <dbl>     <dbl> <chr>         <ch
## 1   -0.657      5.95      6.61     -2.26  0.0246       239    -1.23   -0.0847 Two Sample t~ two
```

We conducted an independent samples t-test to examine if there was a difference in time spent on the internet between males and females. Our t-test showed that males (M = 5.95, SD = 1.95) spent less time on the internet compared to females (M = 6.61, SD = 2.17); $t(239) = -2.26$, p = .025. We are 95% certain that females have values of -1.23 and -0.08 less than males; indicating less time spent on the internet.