# MTH 142: R Assignment 4

name

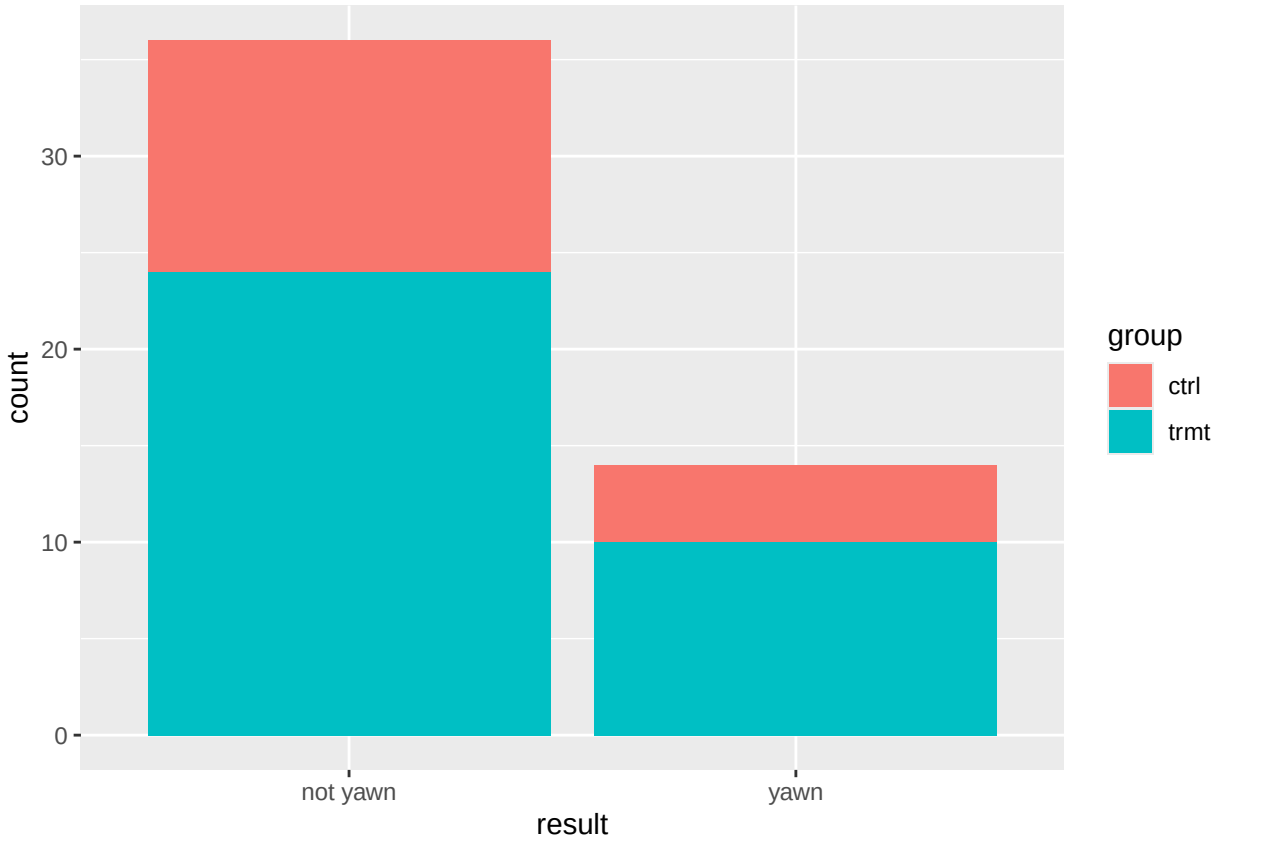
2024-11-15

In 2005 the show "Myth Busters" did an experiment to see if yawning is contagious. In this 4th R assignment you will use data provided in the openintro package to answer that question for yourselves. We will see if the proportions of people that yawn in the treatment group is different then the proportion of people who yawn in the control group. There is a 2 minute video on canvas of that episode. In the show they did not get into statistics, but you will. You will carry out a hypothesis test similar to what you've been doing in the homework, except you will write out the test completely. This assignment is similar to R_Assignment_3, and the instructions for each part have been shortened. In the video on canvas I do this assignment for a different data set, feel free to check that out.

## An exploratory data analysis.

Any graph that shows the number of people who yawned here should be accepted.

```
ggplot(data = yawn, aes(x=result, fill= group))+
  geom_bar()
```



## Check the conditions.*

In this part students have to check success/failure condition and independence/ large sample size for both groups.

Success/faliure: It seems we don't have enough people in the control group to satisfy the conditions of this test. If we treat yawning as a success then we only have 4 successes in the control group.

Independence/ large sample sizes: Our sample is also small, for the control group. It only has 16 people in it.

It would be good to note that we should stop here, but I (the instructor) have asked you to continue the test. Something like: "I should stop here, this test doesn't meet the conditions, but I'm carrying on anyway."

```
count(yawn, result, group)
```

```
## # A tibble: 4 x 3
##   result   group     n
##   <fct>    <fct> <int>
## 1 not yawn ctrl     12
## 2 not yawn trmt     24
## 3 yawn     ctrl      4
## 4 yawn     trmt     10
```

## Write out the hypothesis notation. Include a significance level.

The notation should look something like it does below. In the null we are assuming there are two proportions that are the same. in the alternative we are checking to see if they are different. Alpha should be 0.05, but can be any number less than 0.1. Alpha should not be 0.5 (not that this would be the 50% signifigance level, this is a common mistake students make).

$$H_o : p_{trmt} = p_{ctrl} H_a : p_{trmt} \neq p_{ctrl} \alpha = 0.05$$

## Do the test and state your p-value.

Use prop.test() to have R calculate the p-value for you. State your p-value. note: prop.test() needs four arguments for this test.

- x = number of successes in both groups. Enter as: c(success group 1, successes group 2)
- n = total sample size of both groups. Enter as: c(group1, group2)
- alternative = the alternative hypothesis (I've set this for you).
- correct = FALSE (I've done this for you)

```
# This line of code turns off scientific notation, you don't need to do anything with it.
options(scipen = 100)

prop.test(x = c(4,10) , n = c(16,34) , alternative = "two.sided", correct = FALSE)
```

```
## Warning in prop.test(x = c(4, 10), n = c(16, 34), alternative = "two.sided", :
## Chi-squared approximation may be incorrect
```

```
##
##  2-sample test for equality of proportions without continuity correction
##
## data:  c(4, 10) out of c(16, 34)
## X-squared = 0.10504, df = 1, p-value = 0.7459
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.3057930  0.2175577
```

```
## sample estimates:
##    prop 1    prop 2
## 0.2500000 0.2941176
```

### Write the Conclusion.

The conclusion does not have to be long. But should say something like:

With a pvalue of 0.7459 the null hypothesis should not be rejected. This data provides no evidence that yawning is contagious. Also the results of this test may be misleading, as the conditions for the test were not met.

### Discuss which testing error (I or II) may be possible.

We did not reject the null hypothesis. Its possible that we made a type 2 error.