Common errors

Student name

1/16/2024

Prerequisite

Please read the `voted_na.csv` dataset and assign it as `data` object.

```
data <- read.csv("data/voted_na.csv")
```

Check every step you do; don't run a big block of codes at once

Let's say, I create a new variable called `HispanOld` that takes the value of `TRUE` if a respondent is **Hispanic** (*racename*) and has an **age** equal or above of 60 years. And I want to know if there are differential **gender** (*female*) effects on **voting** across respondents who are `HispanOld` or not:

```
data$HispanOld <- ifelse(data$racename == "Hispanic" & data$age >= 60, TRUE, FALSE)
hoT.data <- subset(data, HispanOld == 1 & female == 1)
hoC.data <- subset(data, HispanOld == 1 & female == 0)
nhoT.data <- subset(data, HispanOld == 0 & female == 1)
nhoC.data <- subset(data, HispanOld == 0 & female == 0)
ho_diff <- mean(hoT.data$voted, na.rm=T) - mean(hoC.data$voted, na.rm=T)
nho_diff <- mean(nhoT.data$voted, na.rm=T) - mean(nhoC.data$voted, na.rm=T)

print(ho_diff)
```

```
## [1] -0.096
```

```
print(nho_diff)
```

```
## [1] 0.02267709
```

Why is there an error? Why is there a `NA`? It is crucial to isolate at which exact step that the error or unexpected result first appears.

Re-estimate the differences in means (`ho_diff` and `nho_diff`) using `tapply()`.

Recall that `tapply()` requires at least three arguments: a **numeric** vector, a vector containing the **index or groups**, and the **function** (`FUN`, e.g. `sum`) to apply in each group.

Hint: in some argument, you will need to use the function `list()`.

```
## Calculate the average first difference by treatment group and by whiteConserv statu
first_diffs <- tapply(data$voted, list(data$female, data$HispanOld), mean, na.rm = TRUE)

first_diffs[2,] - first_diffs[1,]
```

```
##       FALSE        TRUE
##  0.02267709 -0.09600000
```

Knit to `PDF` or ask for help in Slack.

If you successfully compute all the expected quantities, **knit** the document into a `PDF`. If you encounter any issues, create a **minimal reproducible example** by copying the minimum necessary code to reproduce the error or the part where you need help into an `R script` file. Use the `write.csv()` function to save the data file, ensuring to include the correct extension `.csv`.

Once you have either the `PDF` or the **minimal reproducible example**, send me those files via **Slack** in a **private channel** (to me, Ramses). Alongside these files, include either a PDF with the solutions or an R script file + the saved data required for replication. Additionally, mention in the Slack message which **section** you are in (AC or AD).