STA 9750 Project

Presentation:

1. **Description of Data**
    a. Real problem: does social media affect mental health?
    b. Cleaning of data:
        i. Reformat timestamp column -

```
v <- format(as.POSIXct(v,format='%m/%d/%Y
%H:%M:%S'),format='%m/%d/%Y')

v <- c("9/21/2011 0:00:00",  "9/25/2011 0:00:00",  "10/2/2011
0:00:00",
      "9/28/2011 0:00:00",  "9/27/2011 0:00:00")
```

        ii. Removal of:
            1. Column B, E, X, Y, AE, AF, BJ, BK, BL, BM, BN, BO, BP, BQ, BR, BS, BU, BV, BW, AW
        iii. Rename columns:
            1. Code: colnames(data)[2] <- "new_column_name"
        iv. Missing values:
            1. Education and Profession, 470 and 756 - code used: which(is.na(mmc2$`25. Education`))
            2. complete.cases used
        v. Transform variables
            1. Factor -> numerical
                a. Example: mutate(Response=case_when(
                .$Response=="Sometimes" ~ 2,
                .$Response=="Almost Always" ~ 4,
                .$Response=="Almost Never" ~ 1,
                .$Response=="Often" ~ 3))
2. **Association Analysis**
    a. <u>Y variable:</u> Do you think your mental wellbeing would be better if you do not use social media? (column T)
    b. <u>X variables:</u>
        i. In the last 30 days, feeling down, depressed or hopeless (column AQ)
        ii. In the past 30 days, In the past 30 days, do you feel lack of companionship. (column AH)
        iii.
        iv. In last 30 days, How many times, I having trouble sleeping for any other reason? (column BT)
        v. How long have you been using social media account? (column F)
        vi. How much time do you spend daily in social media? (column H)
        vii. How many hours of sleep do you get? (

3. **Regression Models**
    a. single regression x variable: How many friends do you have on social media?
4. **Assessments and interpretations of regression models**
5. **Other techniques you tried for this project**
    a. R-package: dplyr
        i. filter() for selecting rows based on their values
        ii. select() for choosing columns based on their names
        iii. mutate() for adding new variables that are functions of existing variables
        iv. summarise() for calculating summary statistics
        v. arrange() for sorting data
6. **Other graphs/visuals:**
    a. Frequency tables
        i. Do you believe social media is a good thing? (column O)
        ii. Does your emotion get influenced by other's posts? (column R)

Accessing the excel file just change the Users to your customer
socialmediadata <- read.csv("/Users/vrindaarora/Desktop/STAOPR9750/mmc2.csv")
data.frame(socialmediadata)
View(socialmediadata)