

MATH216 HW 2
Due: Friday, September 29th at 11:59pm
Please submit an electronic copy of your assignment to:
***alyford@middlebury.edu* using the subject line:**
MATH216 - HW2

The goal of this lab is to explore the data set *Taboo.csv*. This data set contains information from a subset of participants in a study that sought to investigate possible relationships between ‘taboo’ topics and a measure of a participant’s overall happiness. It’s important to note that the participants were selected randomly, and no two participants were from the same immediate family. The variables in this data set are:

- 1) **ID**- The ID of the person in the study. Participant have been given anonymous IDs to protect their identity.
- 2) **Salary**- The yearly gross income for that participant.
- 3) **Gender**- The participant’s self-reported gender.
- 4) **Sex**- The average number of times the participant (claims to) engage in sexual activity per week on average.
- 5) **Alcohol**- The average number of 12oz. alcoholic drinks the participant consumes weekly.
- 6) **Marijuana**- A binary variable. 1 indicates the participant uses marijuana regularly, defined as smoking marijuana at least once a week. 0 indicates the participant never or rarely uses marijuana.
- 7) **Cigarette** A binary variable. 1 indicates the participant smokes at least a pack of cigarettes per week. 0 indicates the participant smokes less than a pack of cigarettes per week.
- 8) **Happiness**- An indicator of the participant’s happiness, on a scale from 1 to 100. These values were obtained through a survey which asked basic questions about the participant’s well-being and current satisfaction about their life.

Based on the data provided, answer the following questions. You should complete this assignment using R Markdown.

- (1) Based solely on the data, is there evidence of a difference between men and women in the mean number of sexual encounters per week?
- (2) Irrespective of your findings in (1), does there appear to be a difference in the distribution of the number of sexual encounters per week for men versus women?
- (3) Construct a graph of sex versus happiness. Comment about what you see.

- (4) Add a third variable, gender, to your graph in (3). In this graph, highlight the happiest male and female with green dots, and the saddest male and female with red dots. Comment about what this graph tells you about the relationship between sex, happiness, and gender.
- (5) Demonstrate the same data in (4) in an identical fashion using a 1x2 grid of scatterplots. Which of the two representations, (4) or (5), do you prefer and why?
- (6) Is there evidence that different genders are more likely to smoke marijuana? How about cigarettes?
- (7) Split the people who drink alcohol into three separate categories— 1) those who never drink alcohol, 2) those who drink between 1-3 drinks per week, and 3) those who drink 4+ drinks per week. Create a scatter plot of salary versus happiness, and color the points based on the aforementioned three categories. Describe what you see in this plot.
- (8) Write a paragraph or two summarizing your overall findings in this lab. Include at least one graph not mentioned in (1)-(7) that you found particularly interesting.