**STA475 Mini Project (Essential oils)**

The goal of this project is to practice the full process of a statistical analysis, which you will have to do with real clients later in the semester. There are a couple of aspects of this course that are likely going to be different from what you have seen in other statistics courses. The first is that you often will not be told what type of statistical methods you need to employ for a project. Even when a client has an idea about what type of model should be fit to their data, you will need to really consider whether that is the best model. The client is usually not a statistician, so they may not have enough knowledge to determine whether a particular method is appropriate. The second aspect is that the focus of this course is as much about communication as it is about statistical analysis. To be able to adequately address a problem, you need to be able to understand what the research question is from the client, and then you will need to be able to relay your results to them in a manner that they can understand.

Jason will serve as the “client” for this particular project representing the research team he supported as a consultant.

**Background:**

This data set comes from a study about whether essential oils (lavender, rosemary, or both) can impact the quality of sleep, levels of anxiety, and/or depression in patients with Coronary Artery Disease (CAD).  There are four conditions with double-blinding: a control condition where only distilled water was inhaled, and three other conditions where lavender, rosemary, or both were inhaled. Measures of sleep quality, anxiety, and depression were collected at baseline, two weeks, and four weeks of daily inhalation.

The researcher wants to know if any of these outcomes were impacted by the inhalation of the essential oils.

**Participant Selection and Group Assignment:**

A total of 120 patients with heart disease were divided into four groups, with 30 patients randomly assigned to each group.

**A diagram of a flowchart

Description automatically generated**

**Assessment Timeline:**

All groups completed sleep quality (the Pittsburgh Sleep Quality Questionnaire (PSQI)), cardiac depression (Cardiac Depression Questionnaire (CDS)), and anxiety questionnaires (Hospital Anxiety and Depression Questionnaire (HADS)) to establish baseline scores for these variables.

Two weeks after the intervention began, participants completed the questionnaires again to assess changes in sleep quality, cardiac depression, and anxiety.

Participants continued the intervention for another two weeks, followed by a final assessment at the end of the month.

**Hypotheses Testing:**

The first hypothesis predicts significant improvements in sleep quality, reduced anxiety, and depression within each intervention group over one month.

The second hypothesis aims to compare the effectiveness of the different essential oil groups.

The third hypothesis evaluates the interactive effects of the essential oils on sleep, depression, and anxiety to determine which group had the most significant impact.

**Intervention Procedure:**

Group 1 received rosemary extract and were instructed to place two drops on a cloth near their pillow for inhalation during sleep for one month.

Group 2 followed the same procedure using lavender extract.

Group 3 received a combination of rosemary and lavender extracts.

Group 4 received a placebo solution without the active properties of the essential oils.

**Data:**

[EssentialOils.xls](https://www.dropbox.com/scl/fi/i0ddqj0a1odn0cruima6m/Essential_oils.xlsx?rlkey=vj7rcq781yvabishlgr1dkamr&dl=0)

**Variables of interest:**

* ID- individual for matching purposes
* Group: 1-4
* Sex: 1=M, 2=F
* HADS, PSQI, CDS

**Directions:**

Write a succinct 2-3 page report (including tables and figures) detailing your analysis. The report must contain the following sections:

1. Introduction - What question are you addressing?
2. Methods - How are you addressing this question? Detail any data manipulation. Describe the variables.
3. Results - What did you observe in your analysis?  You should have both exploratory and modeling components.
4. Discussion - How did you answer the question?  What caveats/concerns do you have?  What are the next steps/recommendations?
5. References (if applicable) - What software was used? Did you use any other references for your analysis/report?

The data for this problem is simple, but you should be creative about how you conduct and visualize this analysis. You must submit your report via Canvas. The report should be in PDF. You should include R, SAS, or SPSS code at the end of the report but this does NOT count toward your page limit.

This is ideally a team project. You may pair up (no more than 2 per group, strongly encouraged) or work individually. If you pair up, please let me know your group members as soon as possible. You will only submit one joint report per group.